

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 16, 2002, 12:53:58 ; Search time 2228 Seconds
(without alignments)
10904.183 Million cell updates/sec

Title: US-09-822-849A-1
Perfect score: 1800
Sequence: 1 acagacagaactgcgcgtt.....ataaaaaagattacatcc 1800

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	789.6	43.9	808	10	BG681012
2	781	43.4	887	10	Z78306
3	775.6	43.1	819	9	AI760329
4	773.2	43.0	814	9	AUI31388
5	738.6	41.0	829	9	AI355567
6	737.4	41.0	773	9	AI651593
7	717.6	39.9	778	10	BG435706
8	680.8	37.8	847	9	AUI19972
9	675.6	37.5	728	9	AM072830
10	670.4	37.2	1080	10	BG028963
11	662.6	36.8	991	10	BE873929
12	657.8	36.5	796	10	BE538520
13	655.4	36.4	748	10	BG113228
14	648	36.0	908	10	BE545112
15	642.8	35.7	729	10	BF798210
16	631.8	35.1	838	9	AI159665
17	630	35.0	726	10	BG258067

C	18	621.4	34.5	863	9	AI159667	AI159667 ue98d11.x
	19	617.2	34.3	749	10	BE885890	BE885890 601507521
	20	602.6	33.5	620	10	BG403149	BG403149 602418977
	21	593.4	33.0	689	10	BF881346	BF881346 PM0-ET020
C	22	591.8	32.9	603	9	AM341128	AM341128 x294f04.x
	23	579.4	32.2	621	9	AV682561	AV682561 AV682561
	24	570	31.7	838	10	BF185652	BF185652 601817215
	25	565	31.4	583	10	BG656081	BG656081 1b36908.y
	26	563.8	31.3	579	10	BG570725	BG570725 602591839
C	27	550	30.6	561	9	BE217917	BE217917 hv31f07.x
	28	547	30.4	555	10	BF061940	BF061940 7k69d01.x
C	29	545	30.3	555	9	AI627511	AI627511 ly80d11.x
C	30	544.4	30.2	696	9	AI765113	AI765113 w148d08.x
C	31	543	30.2	551	9	AA707321	AA707321 z127c09.s
C	32	534	29.7	547	9	AM152664	AM152664 xf77g02.x
	33	533.8	29.7	583	9	AM392295	AM392295 OVO-ST023
	34	532.8	29.6	762	10	BG170260	BG170260 602321786
C	35	531.2	29.5	548	10	BF433363	BF433363 7q69e02.x
C	36	527	29.3	535	9	AM006558	AM006558 wt25c03.x
C	37	526.2	29.2	539	9	AM471380	AM471380 xw62c11.x
C	38	524.6	29.1	577	9	AI969713	AI969713 w270g11.x
	39	522.8	29.0	555	9	AM842572	AM842572 MR2-CN003
	40	522.4	29.0	705	10	BI655350	BI655350 603283384
C	41	522	29.0	560	10	BE672542	BE672542 7b68g11.x
C	42	522	29.0	566	9	AI694527	AI694527 wd72d02.x
	43	522	29.0	625	9	AA152016	AA152016 z149c09.r
	44	508.4	28.2	632	9	AV702950	AV702950 AV702950
	45	503.2	28.0	729	10	BF219148	BF219148 601882186

ALIGNMENTS

RESULT 1
LOCUS BG681012 808 bp mRNA linear EST 01-MAY-2001
DEFINITION 602628831F1 NCI_CGAP_Skn4 Homo sapiens CDNA clone IMAGE:4753735 5',
mRNA sequence.
ACCESSION BG681012 GI:13912409
VERSION BG681012.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 808)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10613 row: n column: 08
High quality sequence stop: 807.
Location/Qualifiers
1. 808
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4753735"
/clone_lib="NCI_CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI_CGAP library."

FEATURES
source
location/Qualifiers
1. 808

BASE COUNT 280 a 160 c 210 g 158 t

ORIGIN

Query Match 43.9%; Score 789.6; DB 10; Length 808;
Best Local Similarity 99.4%; Pred. No. 1.4e-128;
Matches 803; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 573 TTGAGATGTCACACATGTAATGCTGACACGAGAGTAATGAACTAATCGAGTTGAAA 632
Db 1 TTGAGATGTCACACATGTAATGCTGACACGAGAGTAATGAACTAATCGAGTTGAAA 60
QY 633 GTGTAGAACCTTCAGAAAATGAGCTAGCAAAAGAAATGGAACCAAGAAATGGAATTTGAAA 692
Db 61 GTGTAGAACCTTCAGAAAATGAGCTAGCAAAAGAAATGGAACCAAGAAATGGAATTTGAAA 120
QY 693 TTGAGCCAGATAAAGAAATGTAATCCCTTCTCTGGGAAAGAAATGTCAGTCTTAG 752
Db 121 TTGAGCCAGATAAAGAAATGTAATCCCTTCTCTGGGAAAGAAATGTCAGTCTTAG 180
QY 753 ACATGAAAAAGAGTCTGAGGAAAAAGAAAAAGAAATCTGAGCCCAACCTGAGCCTG 812
Db 181 ACATGAAAAAGAGTCTGAGGAAAAAGAAAAAGAAATCTGAGCCCAACCTGAGCCTG 240
QY 813 TGCTCAACCTCAGCCTCAGTCTCAGCCCACTTCAATCCCACTCCCAACAG 872
Db 241 TGCTCAACCTCAGCCTCAGTCTCAGCCCACTTCAATCCCACTCCCAACAG 300
QY 873 TACTCCAGTCCCAAGCCTCCTCAGCCTGAGGATTTGTCATTAGCTGTTTACAGCCAA 932
Db 301 TACTCCAGTCCCAAGCCTCCTCAGCCTGAGGATTTGTCATTAGCTGTTTACAGCCAA 360
QY 933 CACCCCAAGTACTCAGAGCAAGGCAATTTACTACTGAGAGGAAAGATTTCTGTAG 992
Db 361 CACCCCAAGTACTCAGAGCAAGGCAATTTACTACTGAGAGGAAAGATTTCTGTAG 420
QY 993 AGTCTGTAAACTCAGTGAAGTACCAGTAGAGCCAGTCTTGACAGTACATCCAGAGCA 1052
Db 421 AGTCTGTAAACTCAGTGAAGTACCAGTAGAGCCAGTCTTGACAGTACATCCAGAGCA 480
QY 1053 AGAGCAAAACCAAACTAGAGCAGAGTAGAGTGAAGTGAAGTGAATAAACAAGCAAGA 1112
Db 481 AGAGCAAAACCAAACTAGAGCAGAGTAGAGTGAAGTGAAGTGAATAAACAAGCAAGA 540
QY 1113 GTAGAAAGTCGAGCAGTAGCAGTAGAGTTCAGTTCAGCAGTAGCAGCAGTG 1172
Db 541 GTAGAAAGTCGAGCAGTAGCAGTAGAGTTCAGTTCAGCAGTAGCAGCAGTG 600
QY 1173 GAAGTAGTCCAGCAGTGAAGTAGTAGCAGTTCAGTTCAGTTCAGTTCAGTTCAGTA 1232
Db 601 GAAGTAGTCCAGCAGTGAAGTAGTAGCAGTTCAGTTCAGTTCAGTTCAGTTCAGTA 660
QY 1233 CAAGTGGCAGCAGCAGAGATAGTAGCAGTTCAGTTCAGTTCAGTTCAGTTCAGTA 1292
Db 661 CAAGTGGCAGCAGCAGAGATAGTAGCAGTTCAGTTCAGTTCAGTTCAGTTCAGTA 720
QY 1293 GTCCGAGTAGGGGTCCGGGACATTAATAGATAGAAAGCAGAGAGCGGTGATCGG- 1351
Db 721 GTCCGAGTAGGGGTCCGGGACATTAATAGATAGAAAGCAGAGAGCGGTGATCGG 780
QY 1352 AAGAGAGGATCTTCAAGACTAGAAA 1379
Db 781 AAGAGAGGATCTTCAAGACTAGAAA 808

RESULT 2 278306 887 bp mRNA linear EST 28-JUL-1999
LOCUS 278306/c
DEFINITION HS278306 Human fetal brain S. Meier-Ewert Homo sapiens cDNA clone
1.38 (CEPH) 3', mRNA sequence.
ACCESSION 278306
VERSION 278306.1 GI:1495079
KEYWORDS EST.
SOURCE human,
ORGANISM Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 887)
Neri,C., Albanese,V., Lebre,A.S., Holbert,S., Saada,C., Bougueleret
,L., Meier-Ewert,S., Legall,I., Millaudeau,P., Bul,H., Giudicelli
,C., Massart,C., Guillou,S., Gervy,P., Poullier,E., Rigault,P.,
Weissenbach,J., Lennon,G., Chumakov,I., Dausset,J., Lehrach,H.,
Cohen,D. and Cann,H.M.
Survey of CAG/CTG repeats in human cDNAs representing new genes:
candidates for inherited neurological disorders
Hum. Mol. Genet. 5 (7), 1001-1009 (1996)

TITLE

JOURNAL
MEDLINE
96414310

COMMENT

Contact: Neri C.
Fondation Jean Dausset - CEPH
27 Rue Juliette Dodu, 75010 Paris, France
Related sequence: 278305
3'- sequence (lower strand).
Location/Qualifiers

FEATURES

source

1. .887
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="1.38 (CEPH)"
/clone_lib="Human fetal brain S. Meier-Ewert"
/tissue_type="brain"
/dev_stage="fetus"
/note="cDNA library of S. Meier-Ewert, Max Planck
Inst.f.Mol.Genetics, Berlin, FRG"
BASE COUNT 198 a 222 c 167 g 288 t 12 others
ORIGIN

Query Match 43.4%; Score 781; DB 10; Length 887;
Best Local Similarity 96.3%; Pred. No. 4.6e-127;
Matches 856; Conservative 0; Mismatches 27; Indels 6; Gaps 6;

QY 868 ACCAGTACTCCAGTCCAGCCTCCCT-CTCAGCCTGAGGATTTGTCATTAGCTGTTTAC 926
Db 886 ACCAGTACTCCAGTCCAGCCTCCCTCCCTCCTCAGCCTGAGGATTTGTCATTAGCTGTTTAC 827
QY 927 AGCCAAACCCCAAGTACTCAGAGCAAGGCAATTTACTACTGAGAGGAAGATTTTC 986
Db 826 AGCCAAACCCCAAGTACTCAGAGCAAGGCAATTTACTACTGAGAGGAAGATTTTC 767
QY 987 CTGTAGAGTCTGTAAACTCAGTGAAGTACCAGTAGAGCCAGTCTTGACAGTACATCCAG 1046
Db 766 CTGTAGAGTCTGTAAACTCAGTGAAGTACCAGTAGAGCCAGTCTTGACAGTACATCCAG 707
QY 1047 AGAGCAAGCAAAACCAAACTAGAGCAGAGTAGAGTTCAGTTCAGTTCAGTTCAGTA 1106
Db 706 AGAGCAAGCAAAACCAAACTAGAGCAGAGTAGAGTTCAGTTCAGTTCAGTTCAGTA 647
QY 1107 GCAAGAGTAGAAGTGAAGCAGTAGCAGTTCAGTTCAGTTCAGTTCAGTTCAGTA 1166
Db 646 GCAAGAGTAGAAGTGAAGCAGTAGCAGTTCAGTTCAGTTCAGTTCAGTTCAGTA 587
QY 1167 GCAGTGAAGTAGTTCAGCAGCAGTGAAGTAGTAGCAGTCCG-CAGTAGTTCAGTAGCAGC 1225
Db 586 GCAGTGAAGTAGTTCAGCAGCAGTGAAGTAGTAGCAGTCCGAGTAGTCCAGTAGCAGC 527
QY 1226 TCCAGTCAAGTGGCAGCAGCAGAGATAGTAGCAGTTCAGTTCAGTTCAGTTCAGTA 1285
Db 526 TCCAGTCAAGTGGCAGCAGCAGAGATAGTAGCAGTTCAGTTCAGTTCAGTTCAGTA 467
QY 1286 AGTAGAAGTCGGAGTAGGGGTCCGGGACATTAATAGATAGAAAGCAGAGAGCGCTG 1345
Db 466 AGTAGAAGTCGGAGTA-GGGCCGGGGACATTAATAGATAGAAAGCAGAGAGAG-GCGTG 409
QY 1346 GATGGAGAGAGAGGATCTTCAAGCAGTGAAGAGAGTCAAAATCTTCAAAAGGTGCT 1405
Db 408 GATGGAGAGAGAGGATCTTCAAGCAGTGAAGAGAGTCAAAATCTTCAAAAGGTGCT 349
QY 1406 GGTAGTAGAGATA-CAAAAGATCAAAAGATTAAGATTCGGGTCCGACAGAGAGGTC 1464
Db 348 AGTAGTAGAGATNCCCAAGAGTCNAAGAGATTAAGATTTCC-GTCCGACAGAGAGAGGTC 290

QY 1465 TATATCAGAGAGTAGTCGATCAGGCAAAAGATCTTCAAGAGTGAAAGAGCCCGAAATC 1524
|||||
Db 289 TATATCAGAGAGTAGTCGATCAGGCAAAAGATCTTCAAGAGTGAAAGAGCCCGAAATC 230
QY 1525 AGACAGAAAGACAAAGGCGTTAATGGAAGAGCCAGGCTTTCTTAGCCATTCTTGA 1584
|||||
Db 229 AGACAGAAAGACAAAGGCGTTAATGGAAGAGCCAGGCTTTCTTAGCCATTCTTGA 170
QY 1585 GCAGAGATTCTTGATAAAAAGATTACCTTCTTGTAAGAGAGATGCTGCCTTAAG 1644
|||||
Db 169 GCAGAGATTCTTGATAAAAAGATTACCTTCTTGTAAGAGAGATGCTGCCTTAAG 110
QY 1645 AATTCATGTTGTAATAAAATCTTTTGGAATAACAGACTGTTGTTTACCAGACATTCT 1704
|||||
Db 109 AATTCATGTTGTAATAAAATCTTTTGGAATAACAGACTGTTGTTTACCAGACATTCT 50
QY 1705 TGTACTTTTGCATAATTTTGTAAAGATTATTATCAAAATTAATGTGAG 1753
|||||
49 TGTACTTTTGCATAATTTTGTAAAGATTATTATCAAAATTAATGTGAG 1

RESULT 3

AI760329/c 819 bp mRNA linear EST 20-DEC-1999
LOCUS AI760329
DEFINITION wg97g09.x1 NCI_CGAP_Kid1 Homo sapiens cDNA clone IMAGE:2379232 3'
similar to TR:060899 060899 NUCLEAR PROTEIN SDK3 ; contains element
MIR repetitive element ;, mRNA sequence.

ACCESSION AI760329
VERSION AI760329.1 GI:5175996
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 819)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
Rational Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnlnl.gov/bhrp/image/image.html
Insert length: 916 Std Error: 0.00
Seq primer: -40UP from G1bco
High quality sequence stop: 455.
Location/Qualifiers
1. 819

FEATURES

source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2379232"
/clone_lib="NCI_CGAP_Kid11"
/lab_host="DH10B"
/note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker. Site_1: Not I; Site_2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(clonoids 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."

BASE COUNT 189 a 207 c 121 g 300 t 2 others
ORIGIN

Query Match 43.1%; Score 775.6; DB 9; Length 819;
Best Local Similarity 97.3%; Pred. No. 4e-126;
Matches 798; Conservative 0; Mismatches 21; Indels 1; Gaps 1;

QY 980 GATTTCTCTAGAGTCTGTAAACTCACTGAGGTACCACTAGAGCCAGTCTTGACAGTA 1039
|||||
Db 819 GATTTGCTGTAGAGTCTGTAACTCNCNTAAGGTACAGGTAGAGTCACTTGACAGTA 760
QY 1040 CATCCAGAGAGCAAGAGCAAAACCAAACTAGGAGCAGAGTAGAGTGCAGCTAGAAT 1099
|||||
Db 759 CATCCAGAGAGCAAGAGC-ACATCAATCTAGGAGCAGAGTAGAGTGCAGCTAGAAT 701
QY 1100 AAAACAAGCAAGAGTAGAAGTCAAGCAGTAGCAGTCTTAGTACAGTTCAACC 1159
||
Db 700 AACNCAAGCAAGAGTAGAAGTCAAGCAGTAGCAGTCTTAGTACAGTTCAATC 641
QY 1160 AGTACAGCAGTGGAAAGTAGTCCACAGCAGTGAAGTAGTCCAGTCCAGT 1219
|||||
Db 640 AGTACAGCAGTGGAAAGTAGTCCACAGCAGTGAAGTAGTCCAGTCCAGT 581
QY 1220 AGCAGCTCCAGTACCAAGTGGCAGCAGCAGAGAGTAGTACAGTACAGTACT 1279
|||||
Db 580 AGCAGCTACAGTACCAAGTGGCAGCAGCAGAGAGTAGTACAGTACAGTACT 521
QY 1280 AGTACAGTACCAAGTGGCAGTGGGAGTGGGAGCAGTACATATAGATAGAAAGCAGACAG 1339
|||||
Db 520 AGTACAGTACCAAGTGGGAGTGGGAGTGGGAGCAGTACATATAGATAGAAAGCAGACAG 461
QY 1340 GCGGTGATCGGAAGAGAGAGGATCTTACAGGACTAGAAAGAGTCCACAAATCTTCAAA 1399
|||||
Db 460 AGCGTGATCGGAAGAGAGAGGATCTTACAGGACTAGAAAGAGTCCACAAATCTTCAAA 401
QY 1400 GGTGCTGCTAGTAGAGATACAAAGATCAAAAGATTAAGAAATCCCGTCCGACAGAAAG 1459
|||||
Db 400 GGTGCTGCTAGTAGAGATACAAAGATCAAAAGATTAAGAAATCCCGTCCGACAGAAAG 341
QY 1460 AGGTCTATATCAGAGAGTAGTGCATCAGGCAAAAGATCTTCAAGAGTGAAGAGCCGA 1519
|||||
Db 340 AGGTCTATATCAGAGAGTAGTGCATCAGGCAAAAGATCTTCAAGAGTGAAGAGCCGA 281
QY 1520 AATCAGACAGAAAGACAAAGCGCTTAATGGAAGAGCCAGGCTTTCTTAGCCATTCT 1579
|||||
Db 280 AATCAGACAGAAAGACAAAGCGCTTAATGGAAGAGCCAGGCTTTCTTAGCCATTCT 221
QY 1580 TTGCAGCAGAGAGATTCTTGATTAATAAAGGATTACCTTCTTGTAAAGAGATGCTGCC 1639
|||||
Db 220 TTGCAGCAGAGAGATTCTTGATTAATAAAGGATTACCTTCTTGTAAAGAGATGCTGCC 161
QY 1640 TTAAGAAATGCAATGTTGTAATAAAATCTTTTGGAAATACAGACTGTTGTTTACCAGAC 1699
|||||
Db 160 TTAAGAAATGCAATGTTGTAATAAAATCTTTTGGAAATACAGACTGTTGTTTACCAGAC 101
QY 1700 ATTCTGTACTTTTGGCATATATTTGTTAGAGATTATTTATCAAAATTAATGTGAGTTCCA 1759
|||||
Db 100 ATTCTGTACTTTTGGCATATATTTGTTAGAGATTATTTATCAAAATTAATGTGAGTTCCA 41
QY 1760 AATATGTAAATAATGATATATTAATAAAGATTAACATC 1799
|||||
Db 40 AATATGTAAATAATGATATATTAATAAAGATTAACATC 1

RESULT 4
LOCUS AU131388 814 bp mRNA linear EST 24-OCT-2000
DEFINITION AU131388 NT2RP3 Homo sapiens cDNA clone NT2RP3002489 5', mRNA
sequence.

ACCESSION AU131388
VERSION AU131388
KEYWORDS AU131388.1 GI:10991742
SOURCE EST.
ORGANISM human.

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

D	b		771	GTTCTGCCAGTACATCCAGAGACAAGAACCAAAATTGAAGAGCAGAAGTAAGGT	712
O	y		1087	TCCGACTAGAAATAAACAAGCAGAGTAGAAGTCGAAGCAGTAGCAGTCTTAG	1146
D	b		711	TCCGACTAGAAATTTAACCGACGAAGAGTAGACAAGTCGAAGCAGTAGCAGTCTTAG	652
O	y		1147	TAGCAGTCAACCAGTAGCACGC - AGTCGAAGTAGTTCACAGCAGTGGAGTAGTAGCAGTC	1205
D	b		651	TAGCAGTCCAAACCAGTAGCACGAAGTGGAAGTAGTTCACAGCAGTGGAGTAGTAGCAGTC	592
O	y		1206	GCAGTAGTTCAGTAGCAGCCTCCAGTACCAAGTGGCAGCAGCAGCAGAGATAGTAGCAGTA	1265
D	b		591	GCAGTAGTTCAGTAGCAGCCTCCAGTACCAAGTGGCAGCAGCAGCAGAGATAGTAGCAGTA	532
O	y		1266	GCAGTAG	1325
D	b		531	GCAGTAG	472
O	y		1326	GAAAGCACAGAGAGAGCGGTGGATCCGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1385
D	b		471	GAAAGCACAGAGAGAGCGGTGGATCCGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	412
O	y		1386	ACAAATCTTCAAAGAGGTGGTGGTAGTAGAGATACAAAAGCATCAAAGGATAAGAAATCCC	1445
D	b		411	ACAAATCTTCAAAGAGGTGGTGGTAGTAGAGATACAAAAGCATCAAAGGATAAGAAATCCC	352
O	y		1446	GGTCCGCAGAAAGAGGTCTATATCAGAGAGTAGTCGATCAGCCAAAAGATCTTCAAGAA	1505
D	b		351	GGTCCGCAGAAAGAGGTCTATATCAGAGAGTAGTCGATCAGCCAAAAGATCTTCAAGAA	292
O	y		1506	GTCGAAAGAGCCCCGAAAAATCAGACAGCAAGAGCAAAAGCGCTTAATGGAAGAGCCAGCT	1565
D	b		291	GTCGAAAGAGCCCCGAAAAATCAGACAGCAAGAGCAAAAGCGCTTAATGGAAGAGCCAGCT	232
O	y		1566	TTCTTAGCCATCTTTTGACAGCAGAGATTTCTTGATAAAAAAGGATTACCTTTCCCTGTA	1625
D	b		231	TTCTTAGCCATCTTTTGACAGCAGAGATTTCTTGATAAAAAAGGATTACCTTTCCCTGTA	172
O	y		1626	AAGAGATGCTGCCCTTAAGAATTGCATGTTGTAAAAATCTTTTGGAAAAATACAGACTG	1685
D	b		171	AAGAGATGCTGCCCTTAAGAATTGCATGTTGTAAAAATCTTTTGGAAAAATACAGACTG	112
O	y		1686	TTTTGTTTACAGACATCTCTGTAATTTTGGCATAAATTTTGTAGAGTTATTTATCAAAAT	1745
D	b		111	TTTTGTTTACAGACGTTCTTGTAATTTTGGCATAAATTTTGTAGAGTTATTTATCAAAAT	52
O	y		1746	TATGTGAGGTTCCAAAATATGTAAAAATGATATATATAAAAAAGATTAAAC	1796
D	b		51	TATGTGAGGTTCCAAAATATGTAAAAATGATATATATAAAAAAGATTAAAC	1

ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 773)
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabs-remail.nih.gov Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael

R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 1000 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 444.
 Location/Qualifiers
 1..773

```

/organism-"Homo sapiens"
/db_xref-"taxon:9606"
/clone_1b-"IMAGE:2304956"
/clone_1lb-"NCI_CGAP_GC6"
/tissue_type-"pooled germ cell tumors"
/lab_host-"DH10B"
/note-"Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA
from the normalized library NCI_CGAP_GC4 was prepared, and
ss circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneids
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      169 a      193 c      118 g      291 t
ORIGIN

```

Query Match	41.0%;	Score 737.4;	DB 9;	Length 773;
Best Local Similarity	98.2%;	Pred. No. 2e-119;		
Matches 755;	Conservative 0;	Mismatches 13;	Indels 1;	Gaps 1
OY 1032 TGACAGTACATCCAGAGAGCAAGACCAAAACCAAACTAGGAGCAGAGTAGAGGTCGAG				1091
Db 772 TGACAGGACATCCAGAGAGCAAGACCAAAACCAAACTAGGAGCAGAGTTAGAGGTCGAG				713
OY 1092 CTAGAAATAAAACAAGCAAGAGTAGAAGTCGAAGCAGTAGCAGTAGCAGTTCTAGTAGCA				1151
Db 712 CTAGAAT - AAACAAGCAAGAGTAGAAGTCGAAGCAGTAGCAGTAGCAGTTCTAGTAGCA				654
OY 1152 GTTCACACAGTAGACAGCAGTAGTGAAGTAGTTCACAGCAGTGAAGTAGTAGCAGTCGAGTA				1211
Db 653 GTTCACCCAGTAGACAGCAGTGAAGTAGTTCACAGCAGTGAAGTAGTAGCAGTCGAGTA				594
OY 1212 GTTCACAGTAGCAGCCTCCAGTACAAGTGGCAGCAGCAGCAGAGATAGTAGCAGTAGCACTA				1271
Db 593 GTTCACAGTAGCAGCCTCCAGTACAAGTGGCAGCAGCAGCAGAGATAGTAGCAGTAGCACTA				534
OY 1272 GTAGTAGTAGTGAGAGTAGTAAGTCCGAGTAGGGGTCCGGGACATATATAGAGATAGAAAGC				1331
Db 533 GTAGTAGTAGTGAGAGTAGTAAGTCCGAGTAGGGGTCCGGGACATATATAGAGATAGAAAGC				474
OY 1332 ACAGAAGAGCGCTGGATCGGAAGAGAAGGGATCTTCAGGACTAGAAAGAAGTCACAAAT				1391
Db 473 ACAGAAGAGCGCTGGATCGGAAGAGAAGGGATCTTCAGGACTAGAAAGAAGTCACAAAT				414
OY 1392 CTTCAAAAGGTGGTGTAGTAGAGATACAAAAGGATCAAAAGATTAAGAAATTCGCCGTCGG				1451
Db 413 CTTCAAAAGGTGGTGTAGTAGAGATACAAAAGGATCAAAAGATTAAGAAATTCGCCGTCGG				354
OY 1452 ACAGAAAGAGGTCTATATACAGAGAGTAGTCGATCAGGCAAAAAGATCTTCAGAAAGTGAAA				1511
Db 353 ACAGAAAGAGGTCTATATACAGAGAGTAGTCGATCAGGCAAAAAGATCTTCAGAAAGTGAAA				294
OY 1512 GAGCCCGAAAATCAGACAGAAAGCAAAAAGCGTTAATGGAAGAGCCAGGCTTCTTA				1571
Db 293 GAGACCGAAAATCAGACAGAAAGCAAAAAGCGTTAATGGAAGAGCCAGGCTTCTTA				234
OY 1572 GCCATTCTTGACAGCAGAGATTTCTTCATAAAAAGATTACTTCTTGTAAAGAG				1631

Db 233 GCCATTCCTTCGACGACGAGATTTCTTGATATAAAAAAGCATTACCTTTCCTTGTAAGAGG 174

OY 1632 ATGCTGCCCTAAGATTCATGTTGTATAAAAAATCTTTTGGAAAAATACAGACTGTTGTT 1691
|||||

Db 173 ATGCTGCCCTAAGATTCATGTTGTATAAAAAATCTTTTGGAAAAATACAGACTGTTGTT 114
|||||

OY 1692 TACGACAGATCTCTGACTTTTGTGCAATATTTTGTAGAGTTATTTATCAAAATATGTG 1751
|||||

Db 113 TACGACAGATCTCTGACTTTTGTGCAATATTTTGTAGAGTTATTTATCAAAATATGTG 54
|||||

OY 1752 AGGTTCCAAATATGTATAAAAAATGATATAATAAAAAAGATTACATCC 1800
|||||

Db 53 AGGTTCCAAATATGTATAAAAAATGATATAATAAAAAAGATTACATCC 5

RESULT 7
LOCUS BG435706 778 bp mRNA linear EST 14-MAR-2001
DEFINITION 602506806F1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4604280 5',
mRNA sequence.
ACCESSION BG435706
KEYWORDS BG435706.1 GI:13342212
EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 778)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1CM1339 row: k column: 01
High quality sequence stop: 743.

FEATURES
Source
1..778
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4604280"
/clone_lib="NIH_MGC_79"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: placenta; Vector: pDNR-LIB (Clontech);
Site_1: SfiI (ggccgctgccc); Site_2: SfiI (ggccatgagcc
); 5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGGCCATATGCGC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCCGAGCGCCGACATG-dt(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH-MGC
Library."

BASE COUNT 263 a 153 c 205 g 157 t

ORIGIN

Query Match 39.9%; Score 717.6; DB 10; Length 778;
Best Local Similarity 98.4%; Pred. No. 5.8e-116;
Matches 756; Conservative 0; Mismatches 9; Indels 3; Gaps 3;

OY 551 GTTATGATGCTCTAGAGATGTTGGAATGTCAACATGTAATGCTGACGAGAGTA 610
|||||

Db 1 GTTATGATGCTCTAGAGATGTTGGAATGTCAACATGTAATGCTGACGAGAGTA 60
|||||

OY 611 ATGAACTAATCGAGTTGAAGGTAGAACCTTCAGAAATGAAGCTAGCAAGAATTG 670
|||||

Db 61 ATGGAACATAATCGAGTTGAAAGTGTAGAACCTTCAGAAATGAAGCTAGCAAGAATTG 120

OY 671 GAACGAGAAATGGAATTTGAAATTGAGCCAGATTAAGAATGTAATCCCTTCTCTGGG 730
|||||

Db 121 GAACGAGAAATGGAATTTGAAATTGAGCCAGATTAAGAATGTAATCCCTTCTCTGGG 180
|||||

OY 731 AAGGAAATGTCAGTGTCTTACACATGAAAAAGAGTCTGAGAAAAAGAAAAAGAA 790
|||||

Db 181 AAGGAAATGTCAGTGTCTTACACATGAAAAAGAGTCTGAGAAAAAGAAAAAGAA 240
|||||

OY 791 TCTGAGCCCCAACCTGAGCCCTGTGGCTCAACCTCAGCTCAGCTCAGCTTCAAG 850
|||||

Db 241 TCTGAGCCCCAACCTGAGCCCTGTGGCTCAACCTCAGCTCAGCTCAGCTTCAAG 300
|||||

OY 851 CTTCATCCCAAGTCCCAACCACTACTCCAGTCCCAAGCTCCCTCTCAGCTGAGGATTG 910
|||||

Db 301 CTTCATCCCAAGTCCCAACCACTACTCCAGTCCCAAGCTCCCTCTCAGCTGAGGATTG 360
|||||

OY 911 TCATTAGCTGTTTACAGCCCAACCCCAAGTTACTGAGGAGCAAGGGCAATTACTACT 970
|||||

Db 361 TCATTAGCTGTTTACAGCCCAACCCCAAGTTACTGAGGAGCAAGGGCAATTACTACT 420
|||||

OY 971 GAGAGGAGGATTTCTCTGAGAGTCTGTAAACTCAGTGAAGTACAGTACAGCCAGTC 1030
|||||

Db 421 GAGAGGAGGATTTCTCTGAGAGTCTGTAAACTCAGTGAAGTACAGTACAGCCAGTC 480
|||||

OY 1031 TTGACAGTACATCCAGAGAGCAAGCAAAACCAAACTAGAGAGCAAGTAGAGTGA 1090
|||||

Db 481 TTGACAGTACATCCAGAGAGCAAGCAAAACCAAACTAGAGAGCAAGTAGAGTGA 540
|||||

OY 1091 GCTGAAATTAACAAGCAAGAGTAGAAGTGAAGTGAAGTACAGTACAGTCTAGTAGC 1150
|||||

Db 541 GCTGAAATTAACAAGCAAGAGTAGAAGTGAAGTGAAGTGAAGTACAGTACAGTCTAGTAGC 599
|||||

OY 1151 AGTTCAACCAAGTAGCAGCAGTGAAGTAGTTCACAGTGAAGTAGTACAGTGCAGCA-G 1209
|||||

Db 600 AGTTCAACCAAGTAGCAGCAGTGAAGTAGTTCACAGTGAAGTAGTACAGTGCAGCA 659
|||||

OY 1210 TAGTTCCAGTAGCAGCTCCAGTACCAAGTGGCAGCAGC-AGCAGAGATAGCAGTAGCA 1268
|||||

Db 660 TAGTTCCAGTAGCAGCTCCAGTACCAAGTGGCAGCAGCAGCAGAGATAGCAGTAGCA 719
|||||

OY 1269 CTAGTAGTAGTAGTAGAGTAGAAGTCCGAGTAGGGGTCGGGACATA 1316
|||||

Db 720 CTAGTAGTAGTAGTAGAGTAGAAGTCCGAGTAGGGGTCGGGACATA 767
|||||

RESULT 8
LOCUS AU119972 847 bp mRNA linear EST 19-OCT-2000
DEFINITION AU119972 HEMBA1 Homo sapiens cDNA clone HEMBA1007060 5', mRNA
sequence.
ACCESSION AU119972
VERSION AU119972.1 GI:10935207
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 847)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix

Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES
source
1. .847
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEMBA1007060"
/clone_lib="HEMBA1"
/tissue_type="whole embryo, mainly head"
/dev_stage="embryo, 10 weeks"
/note="Vector: PME18SFL3"

BASE COUNT 295 a 125 c 198 g 225 t 4 others
ORIGIN

Query Match 37.8%; Score 680.8; DB 9; Length 847;
Best Local Similarity 99.0%; Pred. No. 1.6e-109;
Matches 685; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

1109 AAGAGTAGAAGTCGAGCAGTAGCAGTAGCAGTTCTAGTAGCAGTTCACACAGTAGCAGC 1168
1 AAGAGTAGAAGTCGAGCAGTAGCAGTAGCAGTTCTAGTAGCAGTTCACACAGTAGCAGC 60
QY 1169 AGTGAAGTAGTTCAGCAGTAGCAGTAGCAGTTCAGCAGTAGTTCACAGTAGCAGCTCC 1228
DB 61 AGTGAAGTAGTTCAGCAGTAGCAGTAGCAGTAGCAGTTCAGCAGTAGTTCACAGTAGCAGCTCC 120
QY 1229 AGTGAAGTAGTTCAGCAGTAGCAGTAGCAGTAGCAGTTCAGCAGTAGTTCAGTAGTAGAGAGT 1288
DB 121 AGTGAAGTAGTTCAGCAGTAGCAGTAGCAGTAGCAGTTCAGCAGTAGTTCAGTAGTAGAGAGT 180
QY 1289 AGAAGTCGAGTAGAGGGGTCGGGACATAATAGAGATAGAAAGCAGAGAGAGCGGTGAT 1348
DB 181 AGAAGTCGAGTAGAGGGGTCGGGACATAATAGAGATAGAAAGCAGAGAGAGCGGTGAT 240
QY 1349 CGGAGAGAGAGGAGTACTTCAAGACTAGAAAGAGTCAACAATCTTCAAAAAGGTGAT 1408
DB 241 CGGAGAGAGAGGAGTACTTCAAGACTAGAAAGAGTCAACAATCTTCAAAAAGGTGAT 300
QY 1409 AGTAGAGATCAAAAAGAGTCAAGATTAAGATTCCTCGAGAGAGAGGTCTATA 1468
DB 301 AGTAGAGATCAAAAAGAGTCAAGATTAAGATTCCTCGAGAGAGAGGTCTATA 360
QY 1469 TCAGAGAGTAGTCGATCAGGCAAAAGATCTTCAAGAGTGAAGAGCCGAAATCAGAC 1528
DB 361 TCAGAGAGTAGTCGATCAGGCAAAAGATCTTCAAGAGTGAAGAGCCGAAATCAGAC 420
QY 1529 AAGAGAGAGAGAGGCGTTAATGAGAGAGCGCTTCTTAGCCATCTTTGCAGAG 1588
DB 421 AAGAGAGAGAGAGGCGTTAATGAGAGAGCGCTTCTTAGCCATCTTTGCAGAG 480
QY 1589 AAGATTTCTGATAAAAAGATTAATCTTCTTGAAGAGAGTCTGCTTAAGAAT 1648
DB 481 AAGATTTCTGATAAAAAGATTAATCTTCTTGAAGAGAGTCTGCTTAAGAAT 540
QY 1649 GCATGTTGTAATAATCTTTTGAATAACAGACTGTTGTTACAGACATTTCTGTA 1708
DB 541 GCATGTTGTAATAATCTTTTGAATAACAGACTGTTGTTACAGACATTTCTGTA 600
QY 1709 CTTTTCATTAATTTTGTAGAGTTATTTATCAAAATTTATGTGAGTTCCAAATATGTA 1768
DB 601 CTTTTCATTAATTTTGTAGAGTTATTTATCAAAATTTATGTGAGTTCCAAATATGTA 660
QY 1769 AAAATGATTAATTAATAAAAAAGATTACATCC 1800
DB 661 AAAATGATTAATTAATAAAAAAGATTACATCC 692

RESULT 9
AM072830/c 728 bp mRNA linear EST 13-OCT-1999
LOCUS AM072830
DEFINITION xa42e12.x1 NCI_CGAP_Sar4 Homo sapiens cDNA clone IMAGE:2569486 3'
similar to TR:060899 060899 NUCLEAR PROTEIN SDK3 ;, mRNA sequence.

ACCESSION AM072830
VERSION AM072830.1 GI:6027828
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 728)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Possible reversed clone: polyT not found
Seq primer: -40UP from Gibco
High quality sequence stop: 412.

FEATURES
source
1. .728
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2569486"
/clone_lib="NCI_CGAP_Sar4"
/tissue_type="five pooled sarcomas, including myxoid liposarcoma, solitary fibrous tumor, malignant fibrous histiocytoma, gastrointestinal stromal tumor, and mesothelioma"
/lab_host="DH10B"
/note="Organ: connective tissue; Vector: PCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies."

BASE COUNT 137 a 207 c 87 g 293 t 4 others
ORIGIN

Query Match 37.5%; Score 675.6; DB 9; Length 728;
Best Local Similarity 97.7%; Pred. No. 1.3e-108;
Matches 715; Conservative 0; Mismatches 13; Indels 4; Gaps 3;

QY 41 GCGCAGCTGCAAGAGATGGAATGCAATATGCCAAATTAATTAATATATAGAACT 100
DB 728 GCGCAGCTGCAAGAGAG--TGGAATGACATTAATGCCAAATTAATTAATATATAGAACT 672
QY 101 AAGACAAAGCCCATTTGTTTATATTCCTGGAAGATGTGCCAGCTACCCAAAACTA 160
DB 671 AAGACAAAGCCCAATTTGTTTATATTCCTGGAAGATGTGCCAGCTACTC-ANAACTA 613
QY 161 ATAGAAGCTCAGAGAGAAAAATGACGCTTATTGTAAGTAGACGATGGAATTTGCA 220
DB 612 ATAGAAGCTCAGAGAGAAAAATGACGCTTATTGTAAGTAGACGATGGAATTTGCA 553
QY 221 GAACAAATTAATTAATGAGGCTAGGCTAGAGACAAATCAATGAAGAAAAAGACAT 280
DB 552 GAACAAATTAATTAATGAGGCTAGGCTAGAGACAAATCAATGAAGAAAAAGACAT 493
QY 281 CAGGTGTCGTAATGAAGAACAGAGCGGAACAAGAGAGGTTAAGTGGCTACGCA 340
DB 492 CAGGTGTCGTAATGAAGAACAGAGCGGAACAAGAGAGGTTAAGTGGCTACGCA 433
QY 341 GAGGAAGAGTTGAGAGAGAGAGTATCAGCAGATGATGTGAATAGAGAGAGAGCA 400
DB 432 GAGGAAGAGTTGAGAGAGAGAGAGAGTATCAGCAGATGATGTGAATAGAGAGAGCA 373
QY 401 GAGGAAGAGAGAGAGAGAGTATGATGTGATGATGAGAGAGAGAGAGAGAGAG 460

|||||
Db 372 GAGGAAGAGAAAGGAATAGCGATTGTCATAGTATGATGACAGAAAGAACAGAGAGAG 313
QY 461 GAAGAACAACAAACAGGAATGAGGTTAAGATGAGAGAGAACTGAGGTAAGGAAAGT 520
Db 312 GAAGAACAACAAACAGGAATGAGGTTAAGATGAGAGAGAACTGAGGTAAGGAAAGT 253
QY 521 GAGAAGCAGCAGGATAGTACGCTGAGAGAGTTATGATGTCAGATGTTAGAGATGTTAGAGAT 580
Db 252 GAGAAGCAGCAGGATAGTACGCTGAGAGAGTTATGATGTCAGATGTTAGAGATGTTAGAGAT 193
QY 581 GTCGAACATGTAATGTCGACCAAGAGGTAATGGAACCTAATCGAGTTGAAGGTAGAA 640
Db 192 GTCGAACATGTAATGTCGACCAAGAGGTAATGGAACCTAATCGAGTTGAAGGTAGAA 133
QY 641 CCTTCAGAAATGAAAGCTAGCAAGAAATGGAACCAAGAAATGGAATTTGAATTTGAGCCA 700
Db 132 CCTTCAGAAATGAAAGCTAGCAAGAAATGGAACCAAGAAATGGAATTTGAATTTGAGCCA 73
QY 701 GATTAAGAAATGTAATCCCTTCTCTGCGAAGAGAAATGTCAGTCTTACATGAA 760
Db 72 GATTAAGAAATGTAATCCCTTCTCTGCGAAGAGAAATGTCAGTCTTACATGAA 13
QY 761 AAGGAGCTGAG 772
Db 12 AAGGAGCTGAG 1
RESULT 10
LOCUS BG028963 1080 bp mRNA linear EST 24-JAN-2001
DEFINITION 602293090F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4387558 5',
mRNA sequence.
ACCESSION BG028963
VERSION BG028963.1 GI:12418058
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1080)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL10072 row: d column: 23
High quality sequence stop: 643.
Location/Qualifiers
1. 1080
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4387558"
/clone_lib="NIH_MGC_86"
/tissue_type="osteosarcoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: bone; Vector: PCMV-SPORT6; Site: 1; NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.533 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH-MGC Library."
BASE COUNT 378 a 177 c 312 g 213 t
ORIGIN
Query Match 37.2%; Score 670.4; DB 10; Length 1080;
Best Local Similarity 91.9%; Pred. NO. 1.1e-107;

Matches 730; Conservative 0; Mismatches 61; Indels 3; Gaps 2;
QY 407 GAGGAAGAGAAATAGCGATTGTCATAGTATGATGACAGAGAAAGAACAGAGAGAGAA 466
Db 1 GAGGAAGAGAAATAGCGATTGTCATAGTATGATGACAGAGAAAGAACAGAGAGAGAA 60
QY 467 CAAAAACAGGAATGAGGTTAAGATGAGAGAGAACTGAGGTAAGGAAAGTGAAG 526
Db 61 CAAAAACAGGAATGAGGTTAAGATGAGAGAGAACTGAGGTAAGGAAAGTGAAG 120
QY 527 CAGCAGATAGTACGCTGAGAGAGTTATGATGTCAGTCTGAGATGTTAGAAATGTCAAA 586
Db 121 CAGCAGATAGTACGCTGAGAGAGTTATGATGTCAGTCTGAGATGTTAGAAATGTCAAA 180
QY 587 CATGTAATGCTGACCAAGAGGTAATGGAACCTAATCGAGTTGAAGGTAGAACCTTCA 646
Db 181 CATGTAATGCTGACCAAGAGGTAATGGAACCTAATCGAGTTGAAGGTAGAACCTTCA 240
QY 647 GAAATGAGCTAGCAAGAAATGGAACCAAGAAATGGAATTTGAATTTGAGCCAGATTA 706
Db 241 GAAATGAGCTAGCAAGAAATGGAACCAAGAAATGGAATTTGAATTTGAGCCAGATTA 300
QY 707 GAATGTAATCCCTTCTCTGCGAAGAGAAATGTCAGTCTTACATGAAAGAGAG 766
Db 301 GAATGTAATCCCTTCTCTGCGAAGAGAAATGTCAGTCTTACATGAAAGAGAG 360
QY 767 TCTGAGGAAAGAGAGAAAGAAATCTGAGCCCAACCTGAGCTGCAACCTCAG 826
Db 361 TCTGAGGAAAGAGAGAAAGAAATCTGAGCCCAACCTGAGCTGCAACCTCAG 420
QY 827 CCTCAGTCTCAGCCCCAGCTTCACTTCAATCCAGTCCCAACCAAGTCTCAGTCCAG 886
Db 421 CCTCAGTCTCAGCCCCAGCTTCACTTCAATCCAGTCCCAACCAAGTCTCAGTCCAG 480
QY 887 CCTCCCTCTCAGCCCTGAGGATTTGTCATTAAGTGTGTTACAGCCCAACCAAGTCTCAG 946
Db 481 CCTCCCTCTCAGCCCTGAGGATTTGTCATTAAGTGTGTTACAGCCCAACCAAGTCTCAG 540
QY 947 CAGAGCAAGGCGATTACTACCTGAGAGAGAAAGATTTCTGTAGAGTCTGTAACCTC 1006
Db 541 CAGAGCAAGGCGATTACTACCTGAGAGAGAAAGATTTCTGTAGAGTCTGTAACCTC 600
QY 1007 ACTGAGGTACCAAGTAGAGCCAGTCTGACAGTACATCCAGAGCAAGCAACCAAA 1066
Db 601 CTGAGGTACCAAGTAGAGCCAGTCTGACAGTACATCCAGAGCAAGCAACCAAA 658
QY 1067 ACTAGGAGCAGAGTAGAGTGAAGTGAATTAACAAGCAAGAGTAGAGTGAAGC 1126
Db 659 AATAGGAGCAGAGTAGAGTGAAGTGAATTAACAAG- GAGAGTTGAAGTGAAGC 717
QY 1127 AGTAGCAGTAGCAGTCTGAGAGTTCACAGTTCAACCAAGTGAAGTGAAGTTCACAGC 1186
Db 718 CAGTAGCGTATGAGTCAAGTAGAGTTCACACCAAGTGAAGCGTGGAAATGTCACAGAG 777
QY 1187 AGTGAAGTAGTAG 1200
Db 778 GGGAGTATTAGAG 791
RESULT 11
LOCUS BE873929 991 bp mRNA linear EST 20-OCT-2000
DEFINITION 601484590F1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:3886822 5',
mRNA sequence.
ACCESSION BE873929
VERSION BE873929.1 GI:10322705
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 991)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.

OY 378 ATGTAGAAATAGAGAAACAGAGAGAGAAAGAAATAGCATTTCTAGTG 437
|||||
Db 241 ATGTAGAAATAGAGAAACAGAGAGAGAAAGAAATAGCATTTCTAGTG 300
OY 438 ATGCAGAGAAAGAAACAGAGAGAGAAAGAAATAGCATTTCTAGTG 497
|||||
Db 301 ATGCAGAGAAAGAAACAGAGAGAGAAAGAAATAGCATTTCTAGTG 360
OY 498 AGGAACTGAGTAAAGGAAAGTGAAGACAGAGATAGTACGCTGAAGAGTATG 557
|||||
Db 361 AGGAACTGAGTAAAGGAAAGTGAAGACAGAGATAGTACGCTGAAGAGTATG 420
OY 558 ATGTCTAGAGATGTTGAGATGTCACAAATGTAATGCTGACGAGAGTATGAAA 617
|||||
Db 421 ATGTCTAGAGATGTTGAGATGTCACAAATGTAATGCTGACGAGAGTATGAAA 480
OY 618 CTATCGAGTTGAAAGTGTAGAACCTTCAGAAATGAAAGCTAGCAAAAGATTGGAAC 677
|||||
Db 481 CTATCGAGTTGAAAGTGTAGAACCTTCAGAAATGAAAGCTAGCAAAAGATTGGAAC 540
OY 678 AAATGGAATTTGAAATGAGCCAGATTAAGAAATGTAATCCCTTCTCC-TGGGAAAGAG 736
|||||
Db 541 AAATGGAATTTGAAATGAGCCAGATTAAGAAATGTAATCCCTTCTCC-TGGGAAAGAG 600
OY 737 AATGTCAGTCTTTAGA-CATGAAAGAGAGTCTGAGGAA-AGAGAAAGAAATCTG 794
|||||
Db 601 AATGTCAGTCTTTAGACCAAGAGAGTCTGAGGAAAGAGAAAGAAATCTG 660
OY 795 AGCCCAACCTGAGCTGTGCTCAACCTCAGCTCAGCTCAGCCAGCTTCAGCTC 854
|||||
Db 661 AGCCCAACCTGAG--CTGTGTTCAAC--TCAGCTCAGCTCAGCCAGATTCAGATTC 715
OY 855 AATCCAGTCCCAACCACTGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTC 914
|||||
Db 716 AATCCAGTCCCAACCACTGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTC 775
OY 915 TAGCTGTTTACAGCCAC 933
|||||
Db 776 AGCCTTTTAAAGCAAAAC 794

RESULT 13

LOCUS BG113228 748 bp mRNA linear EST 30-JAN-2001
DEFINITION 602283108F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4370904 5',
mRNA sequence.

ACCESSION BG113228
VERSION BG113228.1 GI:12606734

TWO WORDS

EST.

human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 748)

NIH-MGC http://mgs.ncl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

AUTHORS

CONTACT: Robert Strausberg, Ph.D.

EMAIL: cgaabs-remail.nih.gov

TISSUE Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov

Plate: LLM10028 row: 0 column: 01

High quality sequence stop: 695.

Location/Qualifiers

1. 748

/organism="Homo sapiens"

/db_xref="taxon:9606"

FEATURES

Source

/clone="IMAGE:4370904"
/clone_id="NIH_MGC_86"
/tissue_type="osteosarcoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: bone; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.533 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."

BASE COUNT 283 a 115 c 189 g 161 t
ORIGIN

Query Match 36.4%; Score 655.4; DB 10; Length 748;
Best Local Similarity 95.3%; Pred. No. 4.5e-105;
Matches 708; Conservative 0; Mismatches 31; Indels 4; Gaps 3;

OY 1056 GCMAAACCAAACTAGAGCAGAGTAGAGTGCAGCTA-GAAATTAACAAGCAGAGT 1114
|||||
Db 1 GCMAAACCAAACTAGAGCAGAGTAGAGTGCAGCTA-GAAATTAACAAGCAGAGT 60
OY 1115 AGAAGTCGAAGCAGTAGCAGTACAGTCTAGTACAGTTCACACAGTACAGCAGTGA 1174
|||||
Db 61 AGAAGTCGAAGCAGTAGCAGTACAGTCTAGTACAGTTCACACAGTACAGCAGTGA 120
OY 1175 AGTAGTCCAGCAGTGAAGTAGTAGCAGTCCAGTAGTTCACAGTACAGCTCCAGTACA 1234
|||||
Db 121 AGTAGTCCAGCAGTGAAGTAGTAGCAGTCCAGTAGTTCACAGTACAGCTCCAGTACA 180
OY 1235 AGTGCAGCAGCAGCAGAGATAGTAGCAGTACAGTACAGTACAGTACAGTACAGT 1294
|||||
Db 181 AGTGCAGCAGCAGCAGAGATAGTAGCAGTACAGTACAGTACAGTACAGTACAGT 240
OY 1295 CGGAGTAGGGGCGGACATATAGAGATAGAAAGCAGAGAGAGCGGTGATCGGAAG 1354
|||||
Db 241 CGGAGTAGGGGCGGACATATAGAGATAGAAAGCAGAGAGAGCGGTGATCGGAAG 300
OY 1355 AGAAGGATACCTCAGCACTAGAAAGAGTCAAAATCTTCAAAAGGTGGTGTAGTACA 1414
|||||
Db 301 AGAAGGATACCTCAGCACTAGAAAGAGTCAAAATCTTCAAAAGGTGGTGTAGTACA 360
OY 1415 GATACAAAGGATCAAGAGATTAAGATTTCCCGTCCGACAGAAAGAGTCTATATCAGAG 1474
|||||
Db 361 GATACAAAGGATCAAGAGATTAAGATTTCCCGTCCGACAGAAAGAGTCTATATCAGAG 420
OY 1475 AGTAGTCATCAGCAAGAGATCTTCAAGAGTGAAGAGCCGAAATCAGACAGAGAAA 1534
|||||
Db 421 AGTAGTCATCAGCAAGAGATCTTCAAGAGTGAAGAGCCGAAATCAGACAGAGAAA 480
OY 1535 GACAAAGAGCGTTAATGGAAGAGCCAGGCTTTCTTAGCCATCTTTGACAGCAGAGATT 1594
|||||
Db 481 GACAAAGAGCGTTAATGGAAGAGCCAGGCTTTCTTAGCCATCTTTGACAGCAGAGATT 540
OY 1595 TCTTGATTAAGAGATTAACCTTTCCCTTGTAAAGAGATGCTGCTTAAGAAATTGCATGT 1654
|||||
Db 541 TCTTGATTAAGAGATTAACCTTTCCCTTGTAAAGAGATGCTGCTTAAGAAATTGCATGT 600
OY 1655 TGTAAAAAATCTTTTGAAGAAATACAGACTGTTGTTTACAGACATCTTGTACTTTT 1714
|||||
Db 601 TGTAAAAAATC-TTTTGAAGAAATACAGACTGTTGTTTACAGACATCTTGTACTTTT- 658
OY 1715 GCATTAATTTGTAGAGTATTATCAAAATATATGTGAGTTCACAAATATGTAAATG 1774
|||||
Db 659 -GCATTAATTTGTAGAGTATTATCAAAATATATGTGAGTTCACAAATATGTACACAT 717
OY 1775 ATATTAATTAAGAAAGATTAACA 1797
|||||
Db 718 GCTATTAATCAAAAGATACCAAAA 740

RESULT 14

LOCUS BE545112 908 bp mRNA linear EST 09-AUG-2000
DEFINITION 601077850F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:3463621 5',

mRNA sequence.
 BE545112
 BE545112.1 GI:9773757
 EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 908)
 NIH-MGC http://mgs.ncl.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgsbbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L1AM8463 row: c column: 14
 High quality sequence stop: 662.
 Location/Qualifiers
 1..908
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3463621"
 /clone_lib="NIH_MGC_12"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /note="Organ: cervix; Vector: PCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.4 kb. Library prepared by Life
 Technologies."
 BASE COUNT 295 a 175 c 270 g 168 t
 ORIGIN
 Query Match 36.0%; Score 648; DB 10; Length 908;
 Best Local Similarity 93.5%; Pred. No. 8.9e-104;
 Matches 754; Conservative 0; Mismatches 40; Indels 12; Gaps 7;
 OY 636 TAGAACCCTTCAGAAATGAAGAGTACGAAAGATTGGAACGAGAAATGGAATTG 695
 DB 1 TAGAACCCTTCAGAAATGAAGAGTACGAAAGATTGGAACGAGAAATGGAATTG 58
 OY 696 AGCCAGATAAAGATGTAATCCCTTCTCTGCGAAGAGATGTCAGTCTTAGACA 755
 DB 59 AGCCAGATAAAGATGTAATCCCTTCTCTGCGAAGAGATGTCAGTCTTAGACA 118
 OY 756 TGGAAAGAGAGTGTAGGAAAAAGAAAGAAATCTGAGCCCCCACTGAGCCTGTGG 815
 DB 119 TGGAAAGAGAGTGTAGGAAAAAGAAAGAAATCTGAGCCCCCACTGAGCCTGTGG 178
 OY 816 CTGAACCTCAGCCTCAGTCTCAGCCCCAGCTTCAATCCAGTCCCAACAGTAC 875
 DB 179 CTGAACCTCAGCCTCAGTCTCAGCCCCAGCTTCAATCCAGTCCCAACAGTAC 238
 OY 876 TCCAGTCCAGCCTCCTCTCAGCCTGAGGATTTGTCAATTAGCTTTTACAGCCAACAC 935
 DB 239 TCCAGTCCAGCCTCCTCTCAGCCTGAGGATTTGTCAATTAGCTTTTACAGCCAACAC 297
 OY 936 CCCAAGTTACTCAGGAGCAAGGGCATTTACTACCTGAGAGGAGATTTTCTGTAGAGT 995
 DB 298 CCCAAGTTACTCAGGAGCAAGGGCATTTACTACCTGAGAGGAGATTTTCTGTAGAGT 357
 OY 996 CTGTAAACTCAGTGTAGGTACAGGAGGAGTCTTGACAGTACATCCAGAGAGCAAGA 1055
 DB 358 CTGTAAACTCAGTGTAGGTACAGGAGGAGTCTTGACAGTACATCCAGAGAGCAAGA 417
 OY 1056 GCAAAACCAAACTAGAGAGCAAGAGTAGAGTGAAGCTAGAAATTAACAACAAGAGTA 1115
 DB 418 GCAAAACCAAACTAGAGAGCAAGAGTAGAGTGAAGCTAGAAATTAACAACAAGAGTA 477

OY 1116 GAAGTCGAGCAGTAGCAGTAGCAGTCTTAGTAGCAGTTCAACCAAGTAGCAGAGTGGA 1175
 DB 478 GAAGTCGAGCAGTAGCAGTAGCAGTCTTAGTAGCAGTTCAACCAAGTAGCAGAGTGGA 537
 OY 1176 GTAGTTCCAGCAGTGTGAAGTAGTAGCAGTCCAGTAGTCCAGTAGCAGTCCAGTACA 1235
 DB 538 GTAGTTCCAGCAGTGTGAAGTAGTAGCAGTCCAGTAGTCCAGTAGCAGTCCAGTACA 597
 OY 1236 GTGCAGCAGCAG---CAGAGATAGTAGCAGTAGCAGTCTAGTAGTAGTAGAGTAGAA 1292
 DB 598 GTGCAGCAGCAGCAGCAGAGATAGTAGCAGTAGCAGTCTAGTAGTAGTAGAGTAGAA 657
 OY 1293 GTCGAGTAGGGG-TCCGGACATTAAGATAGAGTAAGCA-CAGAAGAGCGGTGATCG 1350
 DB 658 GTCGAGTAGGGGCGCGGACCTAGTAGAGATGAAGAACCGGAGAGCGGGTCCG 717
 OY 1351 GAAGAGAGCGATCTTCAAGACTAGAAAGAGTCAACAATCTTCAAAAGTGTGTAG 1410
 DB 718 ACGGAGCAGGCTCTTCAAGATTAGCAAGAGTCAACCAATCTTCAAA---GGTGTAG 773
 OY 1411 TAGAGATACAAAGCATCAAGATA 1436
 DB 774 GAGTGAGATTCAATGCTTAAGGTTA 799

RESULT 15
 BF798210/c 729 bp mRNA linear EST 12-JAN-2001
 LOCUS BF798210
 DEFINITION RC3-C10043-021000-022-h08 C10043 Homo sapiens CDNA, mRNA sequence.
 ACCESION BF798210
 VERSION BF798210.1 GI:12127199
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 COMMENT
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC3&ct2=RC3-C10043-
 021000-022-h08&ct3=2000-10-02&ct4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 699.

FEATURES
 source
 1..729
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="C10043"
 /dev_stage="Adult"
 /note="Organ: colon; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of

profiles into the puc 18 vector. Reverse transcription of

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 16, 2002, 12:50:53 : Search time 3392 Seconds
(without alignments)
11104.884 Million cell updates/sec

Title: US-09-822-849A-1
Perfect score: 1800
Sequence: 1 acagacagaactgcgcgtt.....ataaaaaagattacatcc 1800

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	1787.2	99.3	2956	9	HSSDK3	Y10351 H. sapiens m
2	1787.2	99.3	3122	9	AF112222	AF112222 Homo sapi
3	1777.6	98.8	3278	9	HSMEWA	Y09703 H. sapiens m
4	1730.4	96.1	2369	6	ARI36786	ARI36786 Sequence
5	1646.6	91.5	2617	9	HSU7718	U77718 Human desmo
6	1603.4	89.1	196319	9	CNS01DT4	AL132639 Human chr
7	1597	88.7	11787	9	AF195139	AF195139 Homo sapi
8	1406.6	78.1	2813	10	MMPININ	Y08701 M. musculus
9	1394	77.4	3893	4	CFU77716	U77716 Canis famil
10	1265	70.3	151836	10	AC079389	AC079389 Rattus no
11	1067	59.3	4167	4	BTU77717	U77717 Bos taurus
12	806.8	44.8	1109	4	PIGSENEU	M55701 Pig neutrop
13	470	26.1	62685	2	AC100145	AC100145 Mus muscu
14	468.4	26.0	49360	2	AC109696	AC109696 Rattus no
15	387.2	21.5	2306	5	XLSDK2	Y10350 X. laevis mR
16	385	21.4	2402	5	XLNUCLEAR	Y10389 Xenopus lae
17	339.4	18.9	349	9	HSU59479	U59479 Human neut
18	320	17.8	62685	2	AC100145	AC100145 Mus muscu
19	308.2	17.1	181089	2	AC094069	AC094069 Rattus no
20	197.8	11.0	461	11	G25957	G25957 human STS E
21	166.4	9.2	239	6	AX197800	AX197800 Sequence
22	166.4	9.2	239	6	AX208402	AX208402 Sequence
23	143	7.9	143	11	G54093	G54093 914 Human H
24	139.8	7.8	272545	2	AC090533	AC090533 Mus muscu
25	136.4	7.6	666	8	AF413050S2	AF413051 Zea mays
26	136.4	7.6	2069	23	E10125	E10125 DNA encodin
27	136.4	7.6	22398	5	FRU271723	AJ271723 Fugu rubr
28	136.2	7.6	144505	2	AL662884	AL662884 Homo sapi
29	136	7.6	144505	2	AL662884	AL645746 Mus muscu
30	133.2	7.4	282611	2	AL645746	AF413050 Zea mays
31	132.4	7.4	543	8	AF413050S1	U93872 Kaposi's sa
32	132.4	7.4	133661	14	U93872	AF397467 Ictalurus
33	130.4	7.2	10115	5	AF397467	AC097839 Rattus no
34	130	7.2	98221	2	AC097839	AL606925 Mus muscu
35	129.2	7.2	237258	2	AL606925	M16230 Strongyloce
36	127.8	7.1	14091	3	SUSSMP1	AC020858 Mus muscu
37	127.6	7.1	186222	2	AC020858	G20375 human STS A
38	127.4	7.1	230	11	G20375	AY016024 Takifugu
39	127	7.1	35793	5	AY016024	I66494 Sequence 14
40	125.6	7.0	7218	6	I66494	AC096822 Rattus no
41	125.4	7.0	17033	2	AC096822	AF305694 Kaposi's
42	123.8	6.9	3127	14	AF305694	AC098456 Rattus no
43	123.6	6.9	257302	2	AC098456	AF413054 Zea mays
44	123.2	6.8	693	8	AF413054	AF139019 Cepaea ne
45	122.8	6.8	624	3	AF139019	

ALIGNMENTS

RESULT 1
HSSDK3
LOCUS
DEFINITION H. sapiens mRNA for nuclear protein SDK3, partial.
ACCESSION Y10351
VERSION Y10351.1 GI:3021391
KEYWORDS nuclear protein SDK3.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Brandner, J.M., Reidenbach, S. and Franke, W.W.
TITLE Evidence that 'pinin', reportedly a differentiation-specific desmosomal protein, is actually a widespread nuclear protein
JOURNAL Differentialiation 62 (3), 119-127 (1997)
MEDLINE 98109012
REFERENCE 2 (bases 1 to 2956)
AUTHORS Brandner, J.M.
TITLE Direct Submission
JOURNAL Submitted (06-JAN-1997) J.M. Brandner, German Cancer Research Centre, Cell Biology, Im Neuenheimer Feld 280, 69120 Heidelberg, FRG

FEATURES
source

Location/Qualifiers
1. 2956
/organism-"Homo sapiens"
/db_xref-"taxon:9606"
/clone-"M2"
/cell_line-"HaCat"
/dev_stage-"adult"
<1. 2138
/codon_start-3
/product-"nuclear protein SDK3"
/protein_id-"CAA71377.1"
/db_xref-"GI:3021392"
/translation-"TLOQLKAKESLKNVDENIRKLTGRDNDVPIQARLLALSGP
GGGRGSLLRGFSDSGGPQAKQDLEGAVSRIGERTRRSROESDPEDDVK
KPAQSSVATSKERTRDLDIQDNDEKQKRNRIFFGLMGTLOKFOESTVATER
OKRQEIQKLEVOAEERKQVENERRELFEERRAKQTELRLLEQVELAQLQEEVME
HNAKIKYIRTKRPHLYIPGRMCPATOKLIEESQKRNALFEGRIEFAQINKME
ARPRQSMKEKHQVLRNEQKAEQKVAOREEELLETQNDHNDVEIEAGEEERK
EIAIYHSDAEKQEEBQKQEMEVKMEETEVRESEKQDSQPEEVMVLEVENKX
VIAQEVMEINRVESVEPSENESELEPEMEFEIEPDKEKSLSPKENVSALDMK
ESEKEKESEPEPEVAQPOPSQPOLQOSQPVLOSQPSQPEDLSLAVLOPTP
QVIOBQHLPERKDFPVEPVKLEVEPEVPLTVHPESKSKTTRSRSRGRANKTSK
SRSS
SRSRGRGNRDKRHRSDRKRDTGLERSHKSSKSGSSSDTKGSKDKNSRDRK
RSISESRSGKRSSRSERDRKSDRDKRR"

BASE COUNT 979 a 570 c 789 g 618 t
ORIGIN

Query Match 99.3%; Score 1787.2; DB 9; Length 2956;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1792; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 1 ACAGACAGAACTGCGGCTTTTGGACAGAAAGTGAGCTTGGCAGCTGCAAGAGAATG 60
DB 590 ACAGACAGAACTGCGGCTTTTGGACAGAAAGTGAGCTTGGCAGCTGCAAGAGAATG 649
OY 61 GAATGAACATATGCCAAATAATTAATATATAGAACTAAGACAAAGCCCAATTGTT 120
DB 650 GAATGAACATATGCCAAATAATTAATATATAGAACTAAGACAAAGCCCAATTGTT 709
OY 121 TTATATTCTGGAAGATGTCTCAGCTACCCAAAACTAATAGAGAGTACAGAGAAA 180
DB 710 TTATATTCTGGAAGATGTCTCAGCTACCCAAAACTAATAGAGAGTACAGAGAAA 769
OY 181 AATGAACGCTTATTGAAGGTAGACGCATCGAATTTGCAGAACAAATAATAATGA 240
DB 770 AATGAACGCTTATTGAAGGTAGACGCATCGAATTTGCAGAACAAATAATAATGA 829
OY 241 GGCTAGGCTAGAGACAATCAATGAAGAAAAAGCATCAGGTGTCGTAATGAAGA 300
DB 830 GGCTAGGCTAGAGACAATCAATGAAGAAAAAGCATCAGGTGTCGTAATGAAGA 889
OY 301 ACAGAGGCGGAAACAAGAGAGGTAGGTGCTCAGCGAGAGAGAGAGAGAGAGAGAG 360
DB 890 ACAGAGGCGGAAACAAGAGAGGTAGGTGCTCAGCGAGAGAGAGAGAGAGAGAGAG 949
OY 361 AGTAATCAGACCAATGATGTAGAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 950 AGTAATCAGACCAATGATGTAGAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1009
OY 421 AGCGATTGTTATAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 1010 AGCGATTGTTATAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1069
OY 481 GGAGTTAGATGAG 540
DB 1070 GGAGTTAGATGAG 1129
OY 541 GCCTGAAGAAGTATGATGCTAGAGATGTTGAGAAATGTCAAACATGTAATGCTGA 600
DB 1130 GCCTGAAGAAGTATGATGCTAGAGATGTTGAGAAATGTCAAACATGTAATGCTGA 1189

OY 601 CCAGAGGTAATGAAACTAATCGAGTTGAAAGTGTAGAACCTTCAGAAATGAAGCTAG 660
DB 1190 CCAGAGGTAATGAAACTAATCGAGTTGAAAGTGTAGAACCTTCAGAAATGAAGCTAG 1249
OY 661 CAAGAATTGCAACAGAAATGGAATTTGAAATTGAGCCAGATTAAGAATGTAAATCCCT 720
DB 1250 CAAGAATTGCAACAGAAATGGAATTTGAAATTGAGCCAGATTAAGAATGTAAATCCCT 1309
OY 721 TTCTCTGGGAAAGAAATGTGAGTGTGTTAGACATGAAAAAGAGTCTGAGAAAAAGA 780
DB 1310 TTCTCTGGGAAAGAAATGTGAGTGTGTTAGACATGAAAAAGAGTCTGAGAAAAAGA 1369
OY 781 AGAAAAAGAAATCTGAGCCCAACCTGAGCCTGTGCTCAACCTCAGCTCAGCTCAGCC 840
DB 1370 AGAAAAAGAAATCTGAGCCCAACCTGAGCCTGTGCTCAACCTCAGCTCAGCTCAGCC 1429
OY 841 CCAGCTTCAGCTTCAATCCAGTCCCAACAGTACTCCAGTCCAGCTCCCTCTCAGCC 900
DB 1430 CCAGCTTCAGCTTCAATCCAGTCCCAACAGTACTCCAGTCCAGCTCCCTCTCAGCC 1489
OY 901 TGAGATTGTGATAGTGTGTTTACAGCCCAACCCCAAGTTACTCAGAGCAAGGCA 960
DB 1490 TGAGATTGTGATAGTGTGTTTACAGCCCAACCCCAAGTTACTCAGAGCAAGGCA 1549
OY 961 TTTACTACCTGAGAGAGAGATTTTCTGTAGAGTCTGTAAAACTCAGTACAGTACAGT 1020
DB 1550 TTTACTACCTGAGAGAGAGATTTTCTGTAGAGTCTGTAAAACTCAGTACAGTACAGT 1609
OY 1021 AGAGCCAGTCTTGACAGTACATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 1610 AGAGCCAGTCTTGACAGTACATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1669
OY 1081 TAGAGTGGAGTGAATAATAAACAAGCAAGAGTGAAGTGAAGCAGTAGCAGTAGCAG 1140
DB 1670 TAGAGTGGAGTGAATAATAAACAAGCAAGAGTGAAGTGAAGCAGTAGCAGTAGCAG 1729
OY 1141 TTCTAGTAGCAGTCAACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 1200
DB 1730 TTCTAGTAGCAGTCAACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 1789
OY 1201 CAGTCGAGTAGTCCAGTAGCAGTCCAGTACAGTACAGTACAGTACAGTACAGTACAG 1260
DB 1790 CAGTCGAGTAGTCCAGTAGCAGTCCAGTACAGTACAGTACAGTACAGTACAGTACAG 1849
OY 1261 CAGTAGCAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 1320
DB 1850 CAGTAGCAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 1909
OY 1321 AGATAGAAAGCAGAGAGAGCGGTGATCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
DB 1910 AGATAGAAAGCAGAGAGAGCGGTGATCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1969
OY 1381 AAGTCACAAATCTTCAAAAAGGTGTGTAGTACAGATCAAAAAGGATCAAAAGATAAGA 1440
DB 1970 AAGTCACAAATCTTCAAAAAGGTGTGTAGTACAGATCAAAAAGGATCAAAAGATAAGA 2029
OY 1441 TTCCCGGTCCGACAGAAAGAGTCTATATCAGAGAGTAGTGCATCAGGCAAAAGATCTTC 1500
DB 2030 TTCCCGGTCCGACAGAAAGAGTCTATATCAGAGAGTAGTGCATCAGGCAAAAGATCTTC 2089
OY 1501 AAGAGTGAAG 1560
DB 2090 AAGAGTGAAG 2149
OY 1561 AGCCTTCTTAGCATCTTTGACAGAGAGAGATTTCTGATAAAAAGATTAACCTTCC 1620
DB 2150 AGCCTTCTTAGCATCTTTGACAGAGAGAGATTTCTGATAAAAAGATTAACCTTCC 2209
OY 1621 TTGTAAAGAGAGAGTGCCTTAAGAAATGTCATGTAATAAAATCTTTTGAAGAAATACA 1680
DB 2210 TTGTAAAGAGAGAGTGCCTTAAGAAATGTCATGTAATAAAATCTTTTGAAGAAATACA 2269
OY 1681 GACTGTTTGTTTACAGACATCTTGTACTTTTGCATATTTTGTAAAGATTAATTATC 1740

Db	1290	CAAGAATTGGAACCAAGAAATGGAAATTTGAAATTTGAGCCAGATAAAGAAATGTAATCCCT	1349
OY	721	TTCTCCTGGAAAGAGAAATGTCAGTGCCTTTAGACATGAAAAAGAGTCTGAGCAAAAAAGA	780
Db	1350	TTCTCCTGGAAAGAGAAATGTCAGTGCCTTTAGACATGAAAAAGAGTCTGAGCAAAAAAGA	1409
OY	781	AGAAAAAGAAATCTGAGCCCCCAACCTGAGCCTTGCGCTCAACCTCAGCCTCAGTCTCAGCC	840
Db	1410	AGAAAAAGAAATCTGAGCCCCCAACCTGAGCCTTGCGCTCAACCTCAGCCTCAGTCTCAGCC	1469
OY	841	CCAGCTCAGCTTCAATCCAGTCCCAACCACTACTCCAGTCCAGCCTCCCTCTCAGCC	900
Db	1470	CCAGCTCAGCTTCAATCCAGTCCCAACCACTACTCCAGTCCAGCCTCCCTCTCAGCC	1529
OY	901	TGAGGATTTGTCATTAAGTGTGTTTACAGCCCAACCCCAAGTTACTCAGAGCAAGGGCA	960
Db	1530	TGAGGATTTGTCATTAAGTGTGTTTACAGCCCAACCCCAAGTTACTCAGAGCAAGGGCA	1589
OY	961	TTTACTACCTGAGAGAGGATTTTCTGTAGAGTCTGTAAACTCAGTGAAGTACAGT	1020
Db	1590	TTTACTACCTGAGAGAGGATTTTCTGTAGAGTCTGTAAACTCAGTGAAGTACAGT	1649
OY	1021	AGAGCCAGTCTTGACAGTACATCCAGAGAGCAAGCAAAACCAAACTAGAGCAGAAAG	1080
Db	1650	AGAGCCAGTCTTGACAGTACATCCAGAGAGCAAGCAAAACCAAACTAGAGCAGAAAG	1709
OY	1081	TAGAGGTGAGCTAGAAATTAACAAGCAAGAGTAGAAGTCGAAGCAGTAGCAGTAGCAG	1140
Db	1710	TAGAGGTGAGCTAGAAATTAACAAGCAAGAGTAGAAGTCGAAGCAGTAGCAGTAGCAG	1769
OY	1141	TTCTAGTAGCAGTTCAACCACTAGCAGCAGTGAAGTAGTTCAGCAGTGAAGTAGTAG	1200
Db	1770	TTCTAGTAGCAGTTCAACCACTAGCAGCAGTGAAGTAGTTCAGCAGTGAAGTAGTAG	1829
OY	1201	CAGTCGAGTAGTTCAGTAGCAGCTCCAGTACAGTGGCAGCAGCAGCAGAGATAGTAG	1260
Db	1830	CAGTCGAGTAGTTCAGTAGCAGCTCCAGTACAGTGGCAGCAGCAGCAGAGATAGTAG	1889
OY	1261	CAGTAGCAGTAGTAGTAGTAGTAGAGAGTAGAAGTCGAGTAGAGGGTGGGAGATAATAG	1320
Db	1890	CAGTAGCAGTAGTAGTAGTAGTAGAGAGTAGAAGTCGAGTAGAGGGTGGGAGATAATAG	1949
OY	1321	AGATAGAAAGCACAGAAAGAGCGGTGATCGGAAGAGAAAGGATACTTCAGAGCTAGAAAG	1380
Db	1950	AGATAGAAAGCACAGAAAGAGCGGTGATCGGAAGAGAAAGGATACTTCAGAGCTAGAAAG	2009
OY	1381	AAGTCACAAATCTTCAAAAGGTGGTAGTAGAGATACAAAAGATCAAAAGATAAGAA	1440
Db	2010	AAGTCACAAATCTTCAAAAGGTGGTAGTAGAGATACAAAAGATCAAAAGATAAGAA	2069
OY	1441	TTCCCGGTCCGACAGAAAGAGTCTATATCAGAGAGTAGTCGATCAGGCAAAAGATCTTC	1500
Db	2070	TTCCCGGTCCGACAGAAAGAGTCTATATCAGAGAGTAGTCGATCAGGCAAAAGATCTTC	2129
OY	1501	AAGAAGTGAAGAGCCGAAATCAGACAGAAAGACAAAAGCGTTAATGGAAGAACCC	1560
Db	2130	AAGAAGTGAAGAGCCGAAATCAGACAGAAAGACAAAAGCGTTAATGGAAGAACCC	2189
OY	1561	AGGCTTTCTTAGCCATCTTTGACAGAGAGATTTCTTGATAAAAAGATTAACCTTCC	1620
Db	2190	AGGCTTTCTTAGCCATCTTTGACAGAGAGATTTCTTGATAAAAAGATTAACCTTCC	2249
OY	1621	TTGTAAGAGAGATGCTGCTTAAGAATTCGATGTAATAAAATCTTTTGGAAATATACA	1680
Db	2250	TTGTAAGAGAGATGCTGCTTAAAGAAATTCGATGTAATAAAATCTTTTGGAAATATACA	2309
OY	1681	GACTGTTGTTTACAGACATCTTGACTTTTGCAATAATTTGTAAGAGTTAATTAATC	1740
Db	2310	GACTGTTGTTTACAGACATCTTGACTTTTGCAATAATTTGTAAGAGTTAATTAATC	2369

RESULT 5
HSU77718

LOCUS	HS077718	2617 bp	mRNA	linear	PRI 01-MAR-1997
DEFINITION	Human desmosome associated protein p1in mRNA, complete cds.				
ACCESSION	U77718				
VERSION	U77718.1	GI:1684846			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 2617)				
TITLE	Ouyang, P. and Sugrue, S. P.				
JOURNAL	Characterization of p1in, a novel protein associated with the desmosome-intermediate filament complex				
MEDLINE	J. Cell Biol. 135 (4), 1027-1042 (1996)				
REFERENCE	97081102				
AUTHORS	2 (bases 1 to 2617)				
TITLE	Sugrue, S. P. and Ouyang, P.				
JOURNAL	Direct Submission				
FEATURES	Submitted (08-NOV-1996) Anatomy and Cell Biology, University of Florida, 1600 SW Archer Road, Gainesville, FL 32610, USA				
source	location/Qualifiers				
	1..2617				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/tissue-type="placenta"				
	31..2262				
	/note="desmosome associated protein; phosphoprotein"				
	/codon_start=1				
	/product="p1in"				
	/protein_id="AAB48304.1"				
	/db_xref="GI:1684847"				
	/translation="MAVAVRTLQEQLEKAKESLVNDEIRKLTGRPNVPIQARRLALSGPGGGRGSLRLRGFSDSGPAPAKORDEGASVRLIGERRRESROESDPEDDVYKRPALQSSVATSKERTRDLIQDQNDKRGKQNRRIFFGLMGLQFKFOESTVATERONRROEIEOKLEVOAEERKOVENERRELFERRAKOTELRLQKVELAOLQLEENNEHNAKIIRYIRTKPHLYIPGRMCPATOKLIESQKRNALFDGRRIEFAEOINKEARPRROSMKEKEHQVVRNEEHKAQOEKGVAQREELVETGQNDVEIEAGEEKEKEIGIVHSDAEKEDEEEOKEMEVKMEETEVRESEKQDSQPEEVMVLEMVENYKHVIADQEMETNRYESVEPSENEASKLEPEMEFEIEPDKECKSLSPKENVSALDMEKESDEKEEKESEPEPVAQPOAQSOQLOQOSQSEPOLOPEPAQPOLOQOPOLQLOQCHAVLOSHPSQPEDLSLAVLPPOVTOEHGFLPERKDFPVESVKLLEVYDPVLTVHPESESETNRSRSGRTNRNRTSRSS				

Db	933	ACACAAGCGCGAACAAGAGAGAGGCTAAGCTGCTCAGCGAGAGAGAGAGAGAGTTGGTGAGAC	992
Qy	361	AGGTAATCAGCAATATGATGTAGAAATAGAGCAACGAGAGAGAGAGAGAGAAAGGAAT	420
Db	993	AGCTAACCAAGCAATATGATGTTGAAATAGAGCAACGAGAGAGAGAGAGAGAGAAAGGAAT	1052
Qy	421	AGCGATTGTTCAATAGTATGATGCAGAGAAAGAACGAGAGAGAGAGAGAGAAACAGGAAT	480
Db	1053	AGGATTTGTTCAATAGTATGATGCAGAGAAAGAACGAGAGAGAGAGAGAGAAACAGGAAT	1112
Qy	481	CGAGTTAAGATGAGAGAGAGAAACTGAGTTAAGGAAAGTGAAGACAGCAGATAGTCA	540
Db	1113	CGAGTTAAGATGAGAGAGAGAAACTGAGTTAAGGAAAGTGAAGACAGCAGATAGTCA	1172
Qy	541	GCCTGAAGAGTATGGATGTGCTAGAGATGTTGAGAAATGTCAAAACATGTAATTGCTGA	600
Db	1173	GCCTGAAGAGTATGGATGTGCTAGAGATGTTGAGAAATGTCAAAACATGTAATTGCTGA	1232
Qy	601	CCAGAGGTAATGGAACCTAATCGAGTTGAAAGTGTAGAACCTTCAGAAAATGAAGCTAG	660
Db	1233	CCAGAGGTAATGGAACCTAATCGAGTTGAAAGTGTAGAACCTTCAGAAAATGAAGCTAG	1292
Qy	661	CAGAATTTGGAACCAAGAAATGGAATTTGAAATTTGAGCCAGATTAAGAAATGTAATCCCT	720
Db	1293	CAGAATTTGGAACCAAGAAATGGAATTTGAAATTTGAGCCAGATTAAGAAATGTAATCCCT	1352
Qy	721	TTCTCTGGGAAAGAGATGTCAAGTCTTAGACATGGAAGAGAGTCTGAGGAAAAAGA	780
Db	1353	TTCTCTGGGAAAGAGATGTCAAGTCTTAGACATGGAAGAGAGTCTGAGGAAAAAGA	1412
Qy	781	AGAAAAAGAACTTGAGCCCCAACCTGAGCCTGTGGCTCAACCTCAGCCTCAGTCCAGCC	840
Db	1413	AGAAAAAGAACTTGAGCCCCAACCTGAGCCTGTGGCTCAACCTCAGCCTCAGTCCAGCC	1472
Qy	841	CCAGCTTCAGCTTCAATCCCAAGTC-----	864
Db	1473	CCAGCTTCAGCTTCAATCCCAAGTC-----	1532
Qy	865	-----CCAACAGTACTCCA	879
Db	1533	ACCTCAGCTTCAGTCTCAGCCCCCAGCTTCAGCCTTCAATCCCAAGTGCATGCACTCCA	1592
Qy	880	GTCCCAAGCCTCCCTCTCAGCCTGAGGATTTGTCATTAAGCTTTTACAGCCCAACCCCA	939
Db	1593	GTCCCAAGCCTCCCTCTCAGCCTGAGGATTTGTCATTAAGCTTTTACAGCCCAACCCCA	1652
Qy	940	AGTTACTCAGAGCAAGGCAATTTACTACCTGAGAGAGAGATTTTCTGTAGAGTCTGT	999
Db	1653	AGTTACTCAGAGCAAGGCAATTTACTACCTGAGAGAGAGATTTTCTGTAGAGTCTGT	1712
Qy	1000	AAAACTCACTGAGTACCACTAGAGCCAGTCTTGACAGTACATCCAGAGAGCAAGCA	1059
Db	1713	AAAACTCACTGAGTACCACTAGAGCCAGTCTTGACAGTACATCCAGAGAGCAAGCA	1772
Qy	1060	AACCAAACTAGAGAGCAGAGTAGAGTGCAGCTAGAAATTAACAAGCAAGAGTAGAAG	1119
Db	1773	AACCAAACTAGAGAGCAGAGTAGAGTGCAGCTAGAAATTAACAAGCAAGAGTAGAAG	1832
Qy	1120	TCGAAGCAGTAGCAGTAGCAGTCTAGTAGCAGTTCAACCACTAGCAGCAGTGAAGTAG	1179
Db	1833	TCGAAGCAGTAGCAGTAGCAGTCTAGTAGCAGTTCAACCACTAGCAGCAGTGAAGTAG	1892
Qy	1180	TTCCAGCAGTGGAGTAGTAGCAGTCCAGTAGTTCACAGTACAGCCTCCAGTACAGTGG	1239
Db	1893	TTCCAGCAGTGGAGTAGTAGCAGTCCAGTAGTTCACAGTACAGCCTCCAGTACAGTGG	1952
Qy	1240	CAGCAGCAGCAGAGATAGTAGCAGTAGCAGTACACTAGTAGTAGTAGAGTAGAAGTCGAG	1299
Db	1953	CAGCAGCAGCAGAGATAGTAGCAGTAGCAGTACACTAGTAGTAGTAGAGTAGAAGTCGAG	2012
Qy	1300	TAGGGTGGGGACATAATAGAGATGAAGCACAGAAGAGCGCTGATCGAAGAGAAAG	1359

Db	2013	TACGGGCCCCGGGACATAAATAGATAGAAAGCACAGAAAGGACCGTGGATTCGGAAGAGAAG	2072
OY	1360	GGATACCTTCAGGACTAGAAAGAAGTCACAAATCTTCAAAAAGGTGGTAGTAGAGATAC	1419
Db	2073	GGATACCTTCAGGACTAGAAAGAAGTCACAAATCTTCAAAAAGGTGGTAGTAGAGATAC	2132
OY	1420	AAAAGGATCAAAAGGATAGAAATTCGCCGGTCCGACAGAAAGAGGTCTATATCAGAGAGTAG	1479
Db	2133	AAAAGGATCAAAAGGATAGAAATTCGCCGGTCCGACAGAAAGAGGTCTATATCAGAGAGTAG	2192
OY	1480	TCGATCAGGCAAAAAGATCTTCAAGAAAGTGAAAGAGAGCCCCGAAAAATCAGACAGGAAAGACAA	1539
Db	2193	TCGATCAGGCAAAAAGATCTTCAAGAAAGTGAAAGAGAGACCGAAAAATCAGACAGGAAAGACAA	2252
OY	1540	AAGCGTTAATGGAAGAAAGCCAGGCTTCTTAGCCATTCTTTGCACAGAAAGATTCTTG	1599
Db	2253	AAGCGTTAATGGAAGAAAGCCAGGCTTCTTAGCCATTCTTTGCACAGAAAGATTCTTG	2312
OY	1600	ATAAAAAAGGATTACCTTTCCTTGTAAAGAGGATGCTGCCTTAAGCAATTGCATGTTGTAA	1659
Db	2313	ATGAAAAAGGATTACCTTTCCTTGTAAAGAGGATGCTGCCTTAAGCAATTGCATGTTGTAA	2372
OY	1660	AAAATCTTTTGGAAAAATACAGAGCTGTTGTTTACCAGAGACATTCTTGACTTTTGCATA	1719
Db	2373	AAAATCTTTTGGAAAAATACAGAGCTGTTGTTTACCAGAGACATTCTTGACTTTTGCATA	2432
OY	1720	ATTTTGTAGAGTTATTTATCAAAATTATGTGAGGTTCCAAAATATGTAAAAATGATAAT	1779
Db	2433	ATTTTGTAGAGTTATTTATCAAAATTATGTGAGGTTCCAAAATATGTAAAAATGATAAT	2492
OY	1780	AATATAAAAAAGATTACATCC 1800	
Db	2493	AATATAAAAAAGATTACATCC 2513	

RESULT 6
CNS01D14
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

CNS01D14 196319 bp DNA linear PRI 29-APR-2001
Human chromosome 14 DNA sequence BAC R-407N17 of library RPCI-11
from chromosome 14 of Homo sapiens (Human), complete sequence.
AL132639
AL132639.4 GI:13897473
HTG.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 196319)
Heilig, R., Petit, J.-L., Vico, V., Dasilva, C., Robert, C., Wincker, P.,
Brottier, P., Cattolico, L., Barbe, V., Pelletier, E., Artiguenave, F.,
Levy, M., Eckenberg, R., Bruls, T., deBerardinis, V., Cruaud, C.,
Gyapay, G., Saurin, W. and Weissendbach, J.
Sequencing of the human chromosome 14
unpublished
2 (bases 1 to 196319)
Genoscope.
Direct Submission
Submitted (28-APR-2001) Genoscope - Centre National de Sequencage

```

COMMENT
-----
Ob: Apr 30, 2001 this sequence version replaced gl:12001725.
-----
Genome Center
Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: SeqRef@genoscope.cns.fr
-----

The following BAC sequence is oriented from the T7 to the SP6 end.
Upstream BAC (overlapping the T7 end) : R-545M17
Downstream BAC (overlapping the SP6 end) : R-647020 (AC-AL157791)
-----
Summary Statistics
-----
Assembly program: Phrap; version 2.0
Quality coverage: 6.71x in Q20 bases; sum-of-contigs

```


Db 9396 CGAGCTAGAAATAAACAAGCAAGAGTAGAAGTTCGAGAGCAGTAGACAGTCTTAGT 9455

Oy 1148 AGCAGTTCAACAGTAGACAGCAGTGAAGTAGTTCAGCAGTGAAGTAGTAGCAGTCCG 1207

Db 9456 AGCAGTTCAACAGTAGACAGCAGTGAAGTAGTTCAGCAGTGAAGTAGTAGCAGTCCG 9515

Oy 1208 AGTAGTTCAGTAGCAGTCCAGTACAGTGGCAGCAGCAGCAGAGATAGTAGCAGTAGC 1267

Db 9516 AGTAGTTCAGTAGCAGTCCAGTACAGTGGCAGCAGCAGCAGAGATAGTAGCAGTAGC 9575

Oy 1268 ACTAGTAGTAGTAGAGAGTAGAGTAGAGTCCGAGTGGGCGGACATTAATAGATAGA 1327

Db 9576 ACTAGTAGTAGTAGAGAGTAGAGTAGAGTCCGAGTGGGCGGACATTAATAGATAGA 9635

Oy 1328 AACACAGAGAGCGGTGATCGAAGAGAGCGATCTTACAGACTAGAAAGAGTCCAC 1387

Db 9636 AACACAGAGAGCGGTGATCGAAGAGAGCGATCTTACAGACTAGAAAGAGTCCAC 9695

Oy 1388 AATCTTCAAAAGGTGGTAGTAGAGATCAAAAGGATCAAAAGGATTAATCCCGG 1447

9696 AATCTTCAAAAGGTGGTAGTAGAGATCAAAAGGATCAAAAGGATTAATCCCGG 9755

Oy 1448 TCCGACAGAAAGAGTCTATATCAGAGAGTAGTTCAGCAGCAAAAGATCTTCAAGAGT 1507

Db 9756 TCCGACAGAAAGAGTCTATATCAGAGAGTAGTTCAGCAGCAAAAGATCTTCAAGAGT 9815

Oy 1508 GAAAGAGCCCAAAATCAGACAGAGAAAGAGCGCTTAATGGAAGAGCCAGCTTT 1567

Db 9816 GAAAGAGAGCCCAAAATCAGACAGAGAAAGAGCGCTTAATGGAAGAGCCAGCTTT 9875

Oy 1568 CTTAGCCATCTTTGCACAGAGAGATTTCTTGATAAAAAGGATTACCTTCTTGTA 1627

Db 9876 CTTAGCCATCTTTGCACAGAGAGATTTCTTGATAAAAAGGATTACCTTCTTGTA 9935

Oy 1628 GAGGATGCTGCTTAAGATTTGCATGTTGTAATAAATCTTTTGGAAATACAGACTGT 1687

Db 9936 GAGGATGCTGCTTAAGATTTGCATGTTGTAATAAATCTTTTGGAAATACAGACTGT 9995

Oy 1688 TGTTCACAGACATCTTCTACTTTTGCATATTTTGTAAAGATTATTATCAAAATTA 1747

Db 9996 TGTTCACAGACATCTTCTACTTTTGCATATTTTGTAAAGATTATTATCAAAATTA 10055

Oy 1748 TGTGAGGTTCCAAATATGTAATAATGATATATAAATAAAGATTACATCC 1800

Db 10056 TGTGAGGTTCCAAATATGTAATAATGATATATAAATAAAGATTACATCC 10108

RESULT 8

PININ 2813 bp mRNA linear ROD 27-OCT-1997

DEFINITION M.musculus mRNA for Pinin protein.

ACCESSION Y08701

VERSION Y08701.1 GI:2570048

KEYWORDS Pinin.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 2813)

AUTHORS Ouyang, P., Zhen, Y.Y. and Sugrue, S.P.

TITLE Cloning and analysis of cDNA encoding murine pinin

JOURNAL Gene 197 (1-2), 115-120 (1997)

MEDLINE 97473502

REFERENCE 2 (bases 1 to 2813)

AUTHORS Ouyang, P.

TITLE Direct Submission

JOURNAL Submitted (07-OCT-1996) P. Ouyang, Chung Gung College Medicine & Technology, 259 Wen-Hua 1st Road, Kuei-San, Tau-Yuan, Taiwan, PROC

FEATURES

source 1..2813

/organism="Mus musculus"

/db_xref="taxon:10090"

/dev_stage="17.5 day embryo"

CDS

19..2196

/codon_start=1

/product="Pinin"

/protein_id="CAA69960.1"

/db_xref="GI:2570049"

/db_xref="SPTREMBL:O35691"

/translation="MAVAVALQQLLEKAKSLKNVDENIRKLTGDPNDVPIQGLT

LALSGPGRGRGSLILRGFSDSGGPPAKORDLEGASRLRGERTRESROESDP

EDDDVKKPALQSSVVAATSKERTDLIDQNDNEKGQRNRRIEGLMGTLOKFOEST

VATERQKRQEIQLKLEQAEERKQVENERRELEKAKOVELAQLQ

EEWENHNKIKYIKRTKTPHLYIPGRMCPATQKLIIESQRKMANLFEGRRIEFAEQ

INMEARPRRPSMKEKEHQVNEEOKAEOEGKVAORKEAEVREGEKQDSQPEEVMVLEAV

EEEEKAGIVHSDAEKQEEEOKEVKEEAEVREGEKQDSQPEEVMVLEAV

ESVKHVAEQEVMETNOVESIESENETSKLEPEMEFDEPKCKSLSPKENINS

QVEKESEEKEKEKEPEPEPEVAQPPQPLPQSQPHSQPHSQPHSQPLC

QPEPLPLAVLPQPPVQIQEQNLPERKDFPLESIKLEVSVEPVLYHSENKSKNT

RSRSRGARNTKTSRSRSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS

SSRDSSTSS

RKSKDKSSRPDRKRSISESSRSRSGRHNDRKHRSDRKRDRDTSGLERHSSKSSSD

BASE COUNT 947 a 504 c 744 g 618 t

ORIGIN

Query Match 78.1%; Score 1406.6; DB 10; Length 2813;

Best Local Similarity 88.1%; Pred. No. 1e-262;

Matches 1610; Conservative 0; Mismatches 184; Indels 33; Gaps 6;

Oy 2 CAGACAGAACTGCGCTTTTGAACAGAAAGTTCGCGCAGCTGCAAGAAATGC 61

Db 622 CAGACAGAGCTGCGCTTTTGAACAGAAAGTTCGCGCAGCTGCAAGAAATGC 681

Oy 62 AATGAACATAATGCCAAATAATTAATATATAGAACTAAGACAAAGCCCATTTGTTT 121

Db 682 AATGAACATAATGCCAAATAATTAATATATAGAACTAAGACAAAGCCCATTTGTTT 741

Oy 122 TATATCTCGAAGAAATGTGCTCAGCTACCCAAACTAATAGAAAGTTCAGAGAA 181

Db 742 TATATCTCGAAGAAATGTGCTCAGCTACCCAAACTAATAGAAAGTTCAGAGAA 801

Oy 182 ATGAACGCTTTATTTGAAGGTAGACGATGAAATTTGCAGAACAAATAAATGAG 241

Db 802 ATGAATGCTTTATTTGAAGGTAGACGATGAAATTTGCAGAACAAATAAATGAG 861

Oy 242 GCTAGGCTAGAACAAATCAATGAAGAAAGAGCATCAGGTGCTGTAATGAGAA 301

Db 862 GCTAGGCTAGAACAAATCAATGAAGAAAGAGCATCAGGTGCTGTAATGAGAA 921

Oy 302 CAGAAGCGGAACAAGAGAGGTAAAGTGGCTCAGAGAGAGAGAGTGGAGAGACA 361

Db 922 CAGAAGCGGAACAAGAGAGGTAAAGTGGCTCAGAGAGAGAGAGTGGAGAGACA 981

Oy 362 GGTATATCAGCAATGATGTAGAAATAGAGAGAGAGAGAGAGAGAGAGAAATA 421

Db 982 GGTATATCAGCAATGATGTAGAAATAGAGAGAGAGAGAGAGAGAGAGAAATA 1041

Oy 422 GCGATTGTTATAGTATGACAGAGAAAGAGAGAGAGAGAGAGAGAGAGAAATG 481

Db 1042 GGTATAGTTCATAGTATGACAGAGAAAGAGAGAGAGAGAGAGAGAGAGAAATG 1101

Oy 482 GAGGTTAAGATGAGAGAGAACTGAGTAAGGAAAGTGAAGAGCAGCAGATAGTCA 541

Db 1102 GAGGTTAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTCA 1161

Oy 542 CCTGAAGAAGTATGATGCTGCTAGAGATGCTGCAATGCTCAACATGTAATGCTGAC 601

Db 1162 CCTGAAGAAGTATGATGCTGCTAGAGATGCTGCAACATGTAATGCTGAC 1221

Oy 602 CAGGAGGTATGAAACTAATGAGTTGAAAGTGTAGAACTTCAGAAATGAGCTAGC 661

Db 1222 CAGGAGGTATGAAACTAATGAGTTGAAAGTGTAGAACTTCAGAAATGAGCTAGC 1281

Oy 662 AAAGAAATTGGAACAGAAATGGAATTTGAAATTTGAGCCAGATTAAGAAATGTAATCCCTT 721

Db 1282 AAAGAAATTGGAACAGAGATGGAATTTGATGTTGAGCCAGATTAAGAAATGTAATCCCTT 1341

OY	722	TCCTCGGAAAGAGAATGTCAGTGTCTTAGACATGGAAAAGAGTCT-----GAG	772
Db	1342	TCCTCGGAAAGAGAATATTAATTCAGGAAGTGGAAAAGAGTCTGAGAGAAAAGAG	1401
OY	773	GAAGAAAGAGAAAAGAAATCTGAGCCCCCAACTGAGCCTGTGGCTCAACCTCAGCCTCAG	832
Db	1402	GAAGAAAGAGAAAAGAACTGAGCCCCCAACTGAGCCTGTGGCTCAGCCTCAGCCCCCA	1461
OY	833	TCACAGCCCCAGC-----TTCAGCTTCAATCCAGTCCCAACCAAGTA	874
Db	1462	CTCAGCCCCCTGCCCCAGAGCCAGCCCCCATCTCTCAGCCCCCACTCCAGGCCCTGTA	1521
OY	875	CTCCAGTCCAGCCTCCCTCTCAGCCTGAGGATTTGTCATTAAGCTGTTTACAGCCAACA	934
Db	1522	CTCCAGTCCAGCCTCTCTGTGAGCCTGAGACTTTGCCATTTAGCTGTTTTCAGCCTCCA	1581
OY	935	CCCCAAGTTACTCAGGAGCAAGGGCATTTACTACTGAGAGGAAGGATTTTCTGTAGAG	994
Db	1582	CCTCAGGTTATTCAGGAGCAAGGGAACTTACTACTGAGCGGAAGGATTTTCTGTAGAG	1641
OY	995	TCCTGTAAGAACTCACTAGGAGTACCAGTAGAGCCAGTCTTGACAGTACATCCAGAGAGCAAG	1054
Db	1642	TCTATAAACTCCCTGAGGTGTCTGTAGAGCCAGTCTTGACAGTACATTCAGAAAACAAA	1701
OY	1055	AGCAAAACCAAACTAGAGAGCAGAAAGTAGAGGTGAGCTAGAAATAAACCAAGCAAGAGT	1114
Db	1702	AGCAAGAAATTAAGACTAGAGAGCAGAGAGCAGGGGGCGAGCAAGAAATAAACCAAGCAAGAGT	1761
OY	1115	AGAACTGAAAGCAGTAGCAGTAGTCAAGTCTTAGTAGCAGTTCACCAAGTAGCAGCAGTGA	1174
Db	1762	AGAACTGAAAGCAGTAGCAGTAGTTCAGTAGCAGTTCAGTAGCAGCAGCAGCGGG	1821
OY	1175	AGTAGTCCAGCAGTAGTGAAGTAGTAGCAGTCCGAGTAGTTCAGTAGCAGCAGTCCAGTACA	1234
Db	1822	AGCAGTCCAGCAGTAGTGCAGCAGCAGTAGTTCGAAGCAGTCCAGCAGCAGCAGCAGCAG	1881
OY	1235	AGTGCGAGCAGCAGCAGAGATAGTAGCAGTAGCAGTACACTAGTAGTAGTAGAGTAGAAGT	1294
Db	1882	AGTGCGAGCAGCAGCAGAGATAGTAGTCCCGCCCTGACAGAAAAGGTCTATATCAGAG	1941
OY	1295	CGAGTAGGGGTGCGGGACATTAATAGATAGAAAGCAGAGAGCGCTGATCGAAG	1354
Db	1942	CGAGTAGGGGTGCGGGACATTAACCGAGATAGGAAGCAGAGAGAGCATGGACCGAAG	2001
OY	1355	AGAAGGATACCTTCAGACTAGAAAGAGTCAACAATCTTCAAAAGGTGCTAGTAGTA	1414
Db	2002	CGAAGGATACCTTCAGACTAGAAAGAGTCAACAAGTCTTCAAAAGGTGCTAGTAGTA	2061
OY	1415	GATACAAAAGGATCAAAAGATAGAATTTCCCGGTCCGACAGAAAGAGCTATATCAGAG	1474
Db	2062	GATACAAAAGGATCAAAAGATAGAATTTCCCGGTCCGACAGAAAGAGCTATATCAGAG	2121
OY	1475	AGTAGTCGATCAGGCAAAAGATCTTCAAGAAGTGAAGAAGCCGAAATCAGACAGGAAA	1534
Db	2122	AGTAGTCGATCAGGCAAAAGATCTTCAAGAAGTGAAGAAGCCGAAATCAGACAGGAAA	2181
OY	1535	GACAAAAGCGCTTAATGGAAGAGCCAGCCTTCTTAGCC-ATTCTTGACACAGAGAT	1593
Db	2182	GACAAAAGCGCTTAATGGAAGAGCCAGCCTTCTTAGCCATTCTTGACACAGAGAT	2241
OY	1594	TTCTTGATAAAAAAGGATTACCTTCTTGTAAAGAGGATGCTGCTTAAGAAATGTCATG	1653
Db	2242	GTCCTGATGAGTA--AAAGATTACCTTGTAAAGAGGATGCTGCTTAAGAAATGTCATG	2298
OY	1654	TTGTAATAATCTTTTGGAAATACAGACTGTTGTTTACAGAGACATCTTGACTTTT	1713
Db	2299	TTGT-TAAATCTTTTGGAAATACAGAGCTGTTGTTTACCTGACATCTTGACTTTT	2357
OY	1714	TGCATAATTTTGTAGAGTATTATTAATAAATATATGTGAGGTTCCAAAATATGTAAAT	1773
Db	2358	TTCATAATTTTGTAGAGTATTATTAATAAATATATGTGAGGTTCCAAAATATGT-TAAAT	2416

OY	1774	GATAATATAAAAAAAAAAGATTACATCC	1800
Db	2417	GATAATAATAAAAAAAAAAGATTACATCC	2443
RESULT	9		
LOCUS	CFU77716		
DEFINITION	CFU77716	3893 bp	mRNA linear MAM 01-MAR-1997
			Canis familiaris desmosome associated protein pinin mRNA, complete cds.
ACCESSION	U77716		
VERSION	U77716.1	GI:1684844	
KEYWORDS			
SOURCE			dog.
ORGANISM			Canis familiaris
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
			Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE			
AUTHORS	1 (bases 1 to 3893)		
TITLE	Ouyang, P. and Sugrue, S.P.		
	Characterization of pinin, a novel protein associated with the		
JOURNAL	desmosome-intermediate filament complex		
MEDLINE	Cell Biol. 135 (4), 1027-1042 (1996)		
REFERENCE	97081102		
AUTHORS	2 (bases 1 to 3893)		
TITLE	Sugrue, S.P. and Ouyang, P.		
JOURNAL	Direct Submission		
	Submitted (08-NOV-1996) Anatomy and Cell Biology, University of		
	Florida, 1600 SW Archer Road, Gainesville, FL 32610, USA		
FEATURES			
source			location/Qualifiers
	1..3893		
	/organism="Canis familiaris"		
	/db_xref="taxon:9615"		
	/cell_line="MDCK"		
	/cell_type="epithelial"		
	/tissue_type="kidney"		
	55..2376		
	/note="desmosome associated protein; phosphoprotein"		
CDS			
	/codon_start=1		
	/product="pinin"		
	/protein_id="AAB48303.1"		
	/db_xref="GI:1684845"		
	/translation="MAVAVFRLQLEAKESLKNVDENIRKLTGRDPNDVRPIOARL		
	LALSGPGGGRGLLRGFSDSGGPAPAKQDLEGAVSRLGGERTRRESROESP		
	EDDDVKRPALOSSVAVATSKERTRDLIODNDEKGKORNRIFGLMGTLOKEOS		
	TVATEROKRROEIQKLEVQAEEERKOVENERRELFERRAKOTELRLIBOKVELAQ		
	QEEMNEHNAKIIRTKTPHLFYILGRMCPATQKLIEESQRKMNALFEGRIEFAE		
	OINKMEARPRRROSMKEKEHQVNNEQKAQEVEGVAAQREEELEETGNQNDVEIEEA		
	GEEEEKIIGIVHSDAEKEOEEOEQOMEVKTIEETFEVRESEKODSQPEEVDYLEM		
	LLHVAVKNVIAEOEVMETNOVESVSENETSKLEPMEFEVEPDKECSLPVREN		
	ASALEMENEPKEERESEPQPEPVRIQLPLQPPEPEPELQPEPOQLQPEPOLQPL		
	OLQLPOPPOSOPOLPLPLPLQPOVOQAOSPOAVLQPOVPVSQPETLPLVL		
	QAPVVIQEOGHLLPERKEFPYESVKLTEVTVEPVLVHSDSKTKTTRSRRGRARN		
	KTSKSRSRSSSSSSSSSSSSSSSSGGSSSSSSSSSSSSSTGSRSRDSSTTT		
	SSSESRSRSGRGHNDRDKHRNSVDKKRDTSGLESHKSSKGGSSSDTSGXDNKNSR		
	SDRRRSISESSRSRSGKRSSRSEDRKSDRKDKRR"		
BASE COUNT	1317 a	697 c	911 g 968 t
ORIGIN			
Query Match	77.4%	Score 1394;	DB 4; Length 3893;
Best Local Similarity	85.4%;	Pred. NO. 2.9e-260;	
Matches 1685; Conservative	0;	Mismatches 115;	Indels 172; Gaps 5;
OY	1	ACAGACAGAAGCTCGGCTTTTGGAACAGAGAAGTTGAGCTTGCAGAGTCGCAAGAAGATG	60
Db	660	ACAGACAGAAGCTCGGCTTTTAGAACAGAGAAGCTTGCAGCTTCGCCAGCTCACAAGAAGATG	719
OY	61	GAATGACATAATGCCAAAATAATTAATATATATAGAACTAAGACCAAAGCCCATTGTT	120
Db	720	GAATGACATAATGCCAAGATAATTAATATACATAGAACTAAGACCAAAGCCCATTGTT	779
OY	121	TTATATTCCTGGAAGATGTGTCCAGCTACCCAAAACTAATAGAGAGTACAGAGAAA	180
Db	780	CTATATTCCTGGAAGATGTGTCCAGCTACCCAAAACTAATAGAGAGTACAGAGAAA	839

OY	181	AATGAACGCTTTATTGTGAAGGTAGACGCATCGAATTTGCAGAACCAATAATAATGA	240
Db	840	AATGAATGCTTTATTGTGAAGGTAGACGCATCGAATTTGCAGAACCAATAATAATGA	899
OY	241	GCGTAGCCCTAGAAAGCAATCAATGAAGAAAAAGAGCATCAGGTGGTGGTAATGAAGA	300
Db	900	GCGTAGCCCTAGAAAGCAATCAATGAAGAAAAAGAGCATCAGGTGGTGGTAATGAAGA	959
OY	301	ACAGAAAGCGGGAACAAGAGAGGGTAAAGTGGCTCAGCGAGAGGAAGAGTTGGAGAGAC	360
Db	960	ACAGAAAGCGGGAACAAGAGAGGGTAAAGTGGCTCAGCGAGAGGAAGAGTTGGAGAGAC	1019
OY	361	AGGTAATCAGCACCAATGATGTAGAAATAGAGGAACAGGAGAGGAAGAAAGGAAAT	420
Db	1020	AGGTAATCAGCACCAATGATGTAGAAATAGAGGAACAGGAGAGGAAGAAAGGAAAT	1079
OY	421	AGCGATTGTCATAGTATGATCGAGAAAGAACAGAGGAGGAAGAACAAAAACAGGAAT	480
Db	1080	AGGTAATGTCATAGTATGATCGAGAAAGAACAGAGGAGGAAGAACAAAAACAGGAAT	1139
OY	481	CGAGCTTAAGATGAGAGGAAACTGAGGTAAGGAAAGTGAGAAAGCAGCAGATAGTCA	540
Db	1140	CGAGCTTAAGATGAGAGGAAACTGAGGTAAGGAAAGTGAGAAAGCAGCAGATAGTCA	1199
OY	541	GCCTGAAGAAGTTATGATGTGCTAGAGATGGTTGAGA-----TGTCAAACATGTAAT	594
Db	1200	ACCGAAGAAGTTATGATGTGCTAGAGATGGTTGACACGTAGCTGTCAAAAATGTAAT	1259
OY	595	TGCTGACCAAGAGGTAATGGAACCTAATCGAGTTGAAAGTGAGAACCTTCAGAAAAATGA	654
Db	1260	TGCTGACCAAGAGGTAATGGAACCTAATCGAGTTGAAAGTGAGAACCTTCAGAAAAATGA	1319
OY	655	AGCTAGCAAAAGAAATTGGAACCAAGAAATGGAATTTGAAATTGAGCCAGATAAAGATGTAA	714
Db	1320	AACTAGCAAAAGAAATTGGAACCAAGAAATGGAATTTGAAATTGAGCCAGATAAAGATGTAA	1379
OY	715	ATCCCTTCTCCTGGAAAGAAATGTCAAGTGTAGACATGMAAAGAGTCTGAGGA	774
Db	1380	ATCTCTTCTCCTGGAGAGAAATGCCAGTGTCTAGAAATGMAAATGAGCCTGAGGA	1439
OY	775	AAAAAGAAAAAAGAAATCTGAGCCCCAACCTGAGCCTGTGCTCAACCTCAGCC-----	828
Db	1440	AAAAAGAAAAAAGAAATCTGAGCCCCAACCTGAGCCTGTGCTCAACCTCAGCCCTTGCC	1499
OY	829	-----	828
Db	1500	TCAGCCTGAGCCTGAGCCTGAGCTACAGCCTGAGCCACAGCCTCAGCTACAACTGAGCC	1559
		829	
Db	1560	ACAGTTACAGCCCCAGCTCCAGCTCCAGCTCCAGCCCCAGCCCCAATCCAGTCCACAGCC	1619
OY	829	-----TCAGTCTCAGCCCCAGCTTCAGCT	852
Db	1620	ACAACCTCAGCTTCAGCTTCGGCTCTCTTCGGCTTCAGCCCCAGCCCCAGGTGCAAGCC	1679
OY	853	TCAATCCCAAGTCCCAACCAAGTACTCCAGTCCCAAGCCTCCCTCAGCCTGAGGATTTGTC	912
Db	1680	CCAGTCCCAAGTCCCAACCAAGTACTCCCAAGCCTCCCTCAGCCTGAGGATTTGTC	1739
OY	913	ATTAGCTGTTTACAGCCAAACCCCAAGTTACTCAGGAGCAAGGGCATTTACTACTGA	972
Db	1740	ATTAGCTGTTTACAGGCAACCAAGTTACTCAGGAGCAAGGGCATTTACTACTGA	1799
OY	973	GAGGAAGGATTTTCTGTAGAGTCTGTAAAACTCACTGAGGTAACCAAGTAGAGCCAGTCTT	1032
Db	1800	GAGGAAGGATTTTCTGTGTGAATCTGTAAAACTCACTGAGGTAACCAAGTAGAGCCAGTCTT	1859
OY	1033	GACAGTACATCCAGAGAGCAAGCAAAAACCAAACTAGACAGAGAGTAGAGGTCGAGC	1092
Db	1860	GATAGTACATTCAGATAGCAAAAACCAAAACCAAAACAGAGCAGAGTAGAGGTCGAGC	1919

QY	1093	TAGAAATAAAACCAAGCAGAGTAGAAGTCGAAGCAGTAGCAGTAGCAGTTCTAGTAGCAG	1152
Db	1920	TAGAAATAAAACCAAGCAGAGTAGAAGTCGCAGCAGTAGCAGTAGTAGTTCTAGTAGCAG	1979
QY	1153	TTCAACCAGTAGCAGCAGTCGAACTAGTTCAGCAGTCGAACTAGTAGCAGTCGCAGTAG	1212
Db	1980	CTCAACCAGTAGCAGCAGTCGAACTAGTTCAGCAGTCGAACTAGTAGTAGTCGAACTAG	2039
QY	1213	TTCCAGTAGCAGCTCCAGTACAGTGGCAGCAGCAGCAGATAGTAGCAGTAGCAGTAG	1272
Db	2040	TTCCAGTAGCAGCTCCAGTACAGTGGCAGCAGCAGCAGATAGTAGCAGTAGCAGTAG	2099
QY	1273	TAGTAGTAGTAGAGAGTAGAAGTCGAGTAGGGGTCGGGCATTAATAGATAGAAAGCA	1332
Db	2100	TAGTAGCAGCAGAGTAGAAGTCGAGTAGGGGCGCGGCATTAACAGATAGAAAGCA	2159
QY	1333	CAGAAGAGCGGTGATCGAAGCAAGCGATCTTACAGACTAGAAAGATCACAAATC	1392
Db	2160	CAGAAGAGCGGTGATCGAAGCAAGCGATCTTACAGACTAGAAAGATCACAAATC	2219
QY	1393	TTCAAAAGGTGGTGTAGTAGAGATACAAAGGATCAAAAGATTAAGATTCCCGTCCGA	1452
Db	2220	TTCAAAAGGTGGTGTAGTAGAGATACAAAGGATCAAAAGATTAAGATTCCCGTCCGA	2279
QY	1453	CAGAAGAGGTCTATATCAGAGAGTAGTCGATCAGGCAAAAGATCTCAAGAAGTGAAG	1512
Db	2280	CAGAAGAGGTCTATATCAGAGAGTAGTCGATCAGGCAAAAGATCTCAAGAAGTGAAG	2339
QY	1513	AGCCCGAAATCAGACAGGAAGACAAAGCGCTTAATGGAAGCCAGGCTTCTTAG	1572
Db	2340	AGACCGAAATCAGACAGGAAGACAAAGCGCTTAATGGAAGCCAGGCTTCTTAG	2399
QY	1573	CC-ATTCTTTGCAGCAGAGAGATTCTTGATTA--AAAAGATTACCTTTCCTTGAAGA	1629
Db	2400	CCATTCTTTGCCGCGAGAGATTCTTGATGAGTAAAGGATTACCTTTCCTTGAAGA	2459
QY	1630	GGATGCTGCCCTTAAGAAITGTCATGTGTAAAAATC-TTTTGGAAAAATACAGACTGTT	1688
Db	2460	GGATGCTGCCCTTAAGAAITGTCATGTGTAAAAATC-TTTTGGAAAAATACAGACTGTT	2519
QY	1689	GTTTACCAGACATCTTGTACTTTTTCATTAATTTTGTAGAGTTATTATCAAAATTAT	1748
Db	2520	GTTTACCAGACATCTTGTACTTTTTCATTAATTTTGTAGAGTTATTATCAAAATTAT	2579
QY	1749	GTGAGGTTCCAAATATGTAAAAATGATTAATTAATAAAAAAGATTAAACATCC	1800
Db	2580	GTGAGGTTCCAAATATGTAAAAATTAATTAATTAATAAAAAAGATTAAACATCC	2631

RESULT	10
LOCUS	AC079389
DEFINITION	AC079389 151836 bp DNA linear ROD 09-NOV-2001 Rattus norvegicus chromosome 6 clone RP31-263K14 strain Brown Norway, complete sequence.
ACCESSION	AC079389
VERSION	AC079389.2 GI:16874847
KEYWORDS	HTG.
SOURCE	Norway rat. Rattus norvegicus
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE	1 (bases 1 to 151836)
AUTHORS	Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Ho,S.-L., Idol,J.R., Karlins,E., Laric,P., Lee-Yin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B., Mastello,C., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Prasad,A., Shevchenko,Y., Stantropop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Tsurgoun,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Zhang,L.-H. and Green,E.D.
TITLE	NISC Comparative Sequencing Initiative

Db	3365	TTCTGTGAATCTGTAAAACTCACTGAGGTACCGACAGAGCCAGTCTTGACGGTACATT	3424
QY	1044	CAGAGACGACAGCAAAACCAAAACCTAGGAGCAGAGTAGAGGTGAGCTAGAAATAAAA	1103
Db	3425	CAGAGACGACGATACGAAACCAAACTAGGAGCAGAGTAGAGGTGAGCTAGAAATAGAA	3484
QY	1104	CAAGCAAGAGTAGAAGTTCGAAGCAGTAGCAGTAGCAGTTCTAGTAGCAGTTCAACCACTA	1163
Db	3485	CCAGCAAGAGTAGAAGTTCGAAGCAGTAGCAGTAGTAGTCTAGTAGCAGTTCAACCACTA	3544
QY	1164	GCAGCAGTGGAACTAGTTCACAGCAGTGGAGTAGTAGCAGTCCGAGTAGTTCACATAGCA	1223
Db	3545	GCAGCAGTGGAAAGCAGTTCCAGCAGTAGTGGCAGTAGCAGTAGTCCGACCAAGT -- AGCAGCT	3601
QY	1224	GCTCCAGTACCAAGTGCAGCAGCAGCAGCAGAGATAGTAGCAGTAGCAGTAGTAGTAGTG	1283
Db	3602	CAAGTAGTACCAAGTGCAGCAGCAGTAGAGACAGTAGCAGCAGCAGCAGTAGCAGTAGTG	3661
QY	1284	AGAGTAGAAGTCGAGTAGGGGTCGGGGACATAATAGAGTAGAAAACACAGAAAGCGG	1343
Db	3662	AGAGCAGAAAGTCGAGTAGGGGCCGGGGGCATAAACAGAGATAGAAAACACAGAAAGCGG	3721
QY	1344	TGCAATCGGAAGAAAGGATCTTACAGGACTAGAAAGAAAGTTCACAAATCTTCAAAAAGTG	1403
Db	3722	TGCAATCGGAAGAAAGGATCTTACAGGACTAGAAAGAAAGTTCACAAATCTTCAAAAAGTG	3781
QY	1404	GTCGTAGTAGAGATACAAAAGCA	1426
Db	3782	GTCGTAGTAGAGATGCAAAAAGCA	3804

RESULT 12	PIGSERNEU	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
	PIGSERNEU	1109 bp	mRNA	linear	MAM 07-MAR-1995			
	Pig neutrophil protein	mRNA, 3' end.						
	M55701							
	M55701.1	GI:164672						
	neutrophil protein.							
	Pig neutrophil, cDNA to mRNA.							
	Sus scrofa							
	Eutharyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
	Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.							
	1 (bases 1 to 1109)							
	Bellavite, P., Bazzoni, F., Cassatella, M.A., Hunter, K.J. and							
	Bannister, J.V.							
	Isolation and characterization of a cDNA clone for a novel							
	serine-rich neutrophil protein							
	Biochem. Res. Commun. 170 (2), 915-922 (1990)							
	90343818							
	Location/Qualifiers							
	1. 1109							

```

/db_xref="taxon:9823"
/clone="B+"
/tissue_type="neutrophil"
<1..857
/codon_start=2
/product="neutrophil protein"
/protein_id="AA63449.1"
/db_xref="GI:164673"
/translation="NSSALEMENEPERKEKESEPOPEPMSOPOPLOPOPPOSOPORC
POAQTOAQCSQSOPALOPLOPETIPLAYLOPPPAIQEGHLPPEKEFPVESAP
KLEVTIVEPVLTVHPESKSRKTKTRSRSGRAANKTSKRSRSSSSSSSTSSSCG
SSSSGSSSSSRSSSSSSSDQVAAAARDSSSTSSSESRSRSGRAHNDRKHKRSVD
HKRDASGLERSHKSSKCGSSRDTKGSKDKNSRSDRKRSLSSSRSRGKRSRSERDRR
SDRKDKRR"

```

Query Match	44.88;	Score 806.8;	DB 4;	Length 1109;
Best Local Similarity	88.2%;	Pred. No. 1.7e-146;		
Matches 975; Conservative	0;	Mismatches 82;	Indels 49;	Gaps 7

OY	736	GAATGTCAGTGCCTTTAGACATGAAAAAGAGTCTGAGAAAAAGAAAGAAATCTGA	795
Db	1	GAATTTAGTGTCTAGAAATGAAAAATGAGCTTGAGAAAAAGAAAGAAATCCGA	60
OY	796	GGCCCAACCTGAGCCCTGTGGCTCAACCTCAGCCTCAGTCTCAGCCCCAGC-----	845
Db	61	GGCCCAAGCTGAGCCTATGTCTCAGCCTCAGCCCCCTGCTCAGCCCCAGCCTCCCA	120
OY	846	-----TTGACGTTCAATCCAGTCCCAACAGT	873
Db	121	GGCACAGCCACAGCCACAGGCCAGACCAGGCCAGTGCCAGTCCCAAGCCAGC	180
OY	874	ACTCCAGTCCCAAGCCTCCTCTCAGCCTGAGATTTGTCAATTAGCTGTTTACAGCCAAC	933
Db	181	ACTCCAGCCCCAGCCTCTCTCTCAGCCTGAGACTATCCCTTAGCTGTTTACAGCCACC	240
OY	934	ACCCCAAGTTACTCAGAGCAAGGCGATTACTACTCTGAGAGGAAGATTTCTCTGTAGA	993
Db	241	ACCTCAGGCTATTACAGAGCAAGGCGATTACCACTGAGAGAAAGAGATTCTCTGTAGA	300
OY	994	GTCGTAAAACTCAGGTACAGTACAGCCAGTCTTGACAGTACATCCAGAGCAAA	1053
Db	301	ATCTGCAAACTGACTGAGGTGACAGTAGAGCCAGTCTTGACAGTACATCCAGAGCAAA	360
OY	1054	GAGCAAAACCAAACCTAGSAGCAGAAGTAGAGGTGCGAGCTAGAAATAAAAACAAGAG	1113
Db	361	GAGCAAAACCAAACCTAGSAGCAGAAGTAGAGGTGCGAGCTAGAAATAAAAACAAGAG	420
OY	1114	TAGAAGTCCAAAGCAGTAGCAGTCTCTAGTAGCACTTCAACCAAGTAGCAGCAGTGG	1173
Db	421	CAGAAAGTCGAAAGCAGTAGCAGTCTCTAGTAGCACTTCAACCAAGTAGCAGCAGTGG	480
OY	1174	AAGTAGTCCCAAGCAGTGGAAAGTAGACAGTCCGAGTAGTCTCCAGTCCAGTCCAGTAC	1233
Db	481	AAGCAGTCTAGCAGTAGTGGCAGTAGTAGTAGTTCGAAGTAGTCCAGTAGCAGTCC-GATC	539
OY	1234	AAGTGGCAGCAGCAGC-AGAGATAGTAGCAGTAGCAGTCTAGTAGTAGTAGAGTAGAA	1292
Db	540	AAGTGGCAGCAGCAGCAGTAGAGACAGTAGCAGCAGCAGTCTAGTAGTAGAGTAGAA	599
OY	1293	GTCGAGTAGGGGTCGGGACATATAGAGATAGAAAGCAGAAAGGCGTGCATCCGA	1352
Db	600	GTCGAAGTAGGGGCGGGGCACATACAGAGATAGAAACACAGGAGGTGTGATCCGA	659
OY	1353	AGAGAAGGCACTCTCAGGACTAGAAAGAGTCAACAATCTTCAAAAGGTGGTGTAGTA	1412
Db	660	AGCGAGGAGTCTTCTGACTAGAAAGAGTCAACAATCTTCAAAAGGTGGTGTAGTA	719
OY	1413	GAGATACAAAAGGATCAAAAGGATAAGATTCCCGGTCCGACAGAAAGAGTCTATATCAG	1472
Db	720	GAGATACAAAAGGATCAAAAGGATAAGATTCCCGGTCTGACAGAAAGAGTCTATATCAG	779
OY	1473	AGAGTAGTCATCAGGCCAAAAAGATCTTCAAGAGTGAAGAGAGCCGAAAAATCAGACAGGA	1532
Db	780	AGAGTAGTCATCAGGCCAAAAAGATCTTCAAGAGTGAAGAGAGCCGAAAAATCAGACAGGA	839
OY	1533	AAGACAAAAGCGTTAATGGAAGAGCCAGGCTTTCTTACCATTTCTTGACAGCAGAGA	1592
Db	840	AAGACAAAAGCGTTAATGGAAGAGCCAGGCTTTCTTACCATTTCTTGACAGCAGAGA	899
OY	1593	TTTCTTGATAA--AAAAGATTACCTTCTGTAAAGAGATGCTGCTTAAGAAATGTC	1650
Db	900	TTTCTTGATGAGTAAGAGATTACCTTCTGTAAAGAGATGCTGCTTAAGAAATGTC	959
OY	1651	ATGTTGTAATAAAATCTTTTGGAAAAATACAGACTGTTGTTTACAGACATTTCTGTACT	1710
Db	960	ATGTTGT-AAAAATTTTGGAAAAATACAGACTGTTGTTTACAGAC-TTCTTGTAC-	1016
OY	1711	TTTTCATATTTTGTAAAGAGTTATTTATCAAAATTTATGTAGGTTCCAAATATGTAA	1770
Db	1017	TTTTCATATTTTGTAAAGAGTTATTTATCAAAATTTATGTAGGTTCCAAATATGTAA	1076
OY	1771	AATGATATATATAAAAAAGATTAAAC	1796

Db 8010 AGGATGAGAGCAGACTGGGTAAGAGAGCTGAGAAGCAGACAGATAGTCAGCCAGAA 8069
QY 548 GAAGTATGATGCTGCTAGAGATGTTGAGATGTCACATGTAATGCTGACCAGAG 607
Db 8070 GAAGTATGATGCTGCTAGAGATGTTGAGATGTCACATGTAATGCTGACCAGAG 8129
QY 608 GTAATGAACTAATCGAGTTGAAGTGTGAACCTTCAGAAAATGAACTAGCAAGAA 667
Db 8130 GTGATGAGACTAATCAAGTTGAAGTGTGAACCTTCAGAAAATGAACTAGCAAGAA 8189
QY 668 TTGAACCCAGAAATGGAATTTGAATTTGAGCCAGATAAAGATGTAATCTCTCTCT 727
Db 8190 TTGAGCCAGAGATGGAATTTGAGCTTGAAGCCAGATAAAGATGTAATCTCTCTCT 8249
QY 728 GGGAAAGAGATGTCAGTCTTGAACATGGAAGAGAGTCTGAGGAAAGAGAGAA 785
Db 8250 GGAAGAGAGATGTAATTTCTCTGAGAGTGAAGAGAGTCTGAGGAAAGAGAGAA 8305

SULT 15
SDK2
LOCUS
DEFINITION X.laevis mRNA for nuclear protein SDK2.
ACCESSION Y10350.1 GI:3021599
KEYWORDS nuclear protein SDK2.
SOURCE African clawed frog.
ORGANISM Xenopus laevis

REFERENCE 1 (bases 1 to 2306)
AUTHORS Brandner, J.M., Reidenbach, S. and Franke, W.W.
TITLE Evidence that 'pinin', reportedly a differentiation-specific desmosomal protein, is actually a widespread nuclear protein
JOURNAL Differentialiation 62 (3), 119-127 (1997)
MEDLINE 98109012
REFERENCE 2 (bases 1 to 2306)
AUTHORS Brandner, J.M.
TITLE Direct Submission
JOURNAL Submitted (06-JAN-1997) J.M. Brandner, German Cancer Research Centre, Cell Biology, Im Neuenheimer Feld 280, 69120 Heidelberg, FRG

FEATURES
source
1. 2306
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="M3"
/cell_type="oocyte"
/clone_lib="Uni-ZAP, Strategene"
/dev_stage="adult"
27. 2096
/codon_start=1
/evidence-not_experimental
/product="nuclear protein SDK2"
/protein_id="CAA71376.1"
/db_xref="GI:3021600"
/db_xref="SPTREMBL:O73620"

translation="MAEAVRTLOEQLERAKESQKNVDENIKLITGRDNETRPAQARR
LSITGPNAIGRGRGGLLRGLSDSGPPAKQDPEFEGALNRLGERRQRESROES
DAEDEDKRALQSSVATSKERTRLIDQNMDEKQKORNRMFGLMGTLOKFKO
ESNVATEOLRRHEIEOKLEVOAEERKLYESERELFEERTKQTELRLEOKVELA
HLOEWMNDHKKIKIRTKPHLFTIPGRMCPOSQKFDESOQKLNATFESRLEF
AEOLNKEARPRROSMKDRDQEKPEPEONEHETEDREDKMSRHNREVEYGNQPDVD
MEDTLEEDKVTISIDVEMADARVEDLEDSEKMEKIEEGRVVEESQNSQWNE
REVEVEDEVEKEETKEIKONISEDOFTFEISOEVLDMENLKDNREPSVQAEEN
FTFEATEATEPESEREVESEPERLOKEPDLMFQPPANTESRTQPEHNFVDEEPFE
KEOMAAVITAEQVEYAPKPVASVHTKRSRGRNKNMGKSRSSSSSSSSSSST
SGSS
HKRDGAGVERSHKSSKSGRDSKSGKSSRSRSDRRTTSEGSRSKRPSSNERDRKS
DRKDKRR"

BASE COUNT 861 a 384 c 629 g 432 t
ORIGIN

Query Match 21.5%; Score 387.2; DB 5; Length 2306;
Best Local Similarity 59.2%; Pred. No. 3.7e-65;
Matches 922; Conservative 0; Mismatches 528; Indels 108; Gaps 11;

QY 1 ACAGACAGAACTGCGGCTTTGGACACAGAAAGTTGAGCTTGGCCAGCTGCAAGAAAGATG 60
Db 638 ACAGACAGAGCTAGCTCTTCTAGAACAGAGAGCTGAGCTTGCACACCTGCAAGAGAGTG 697
QY 61 GAATGAACATATATGCCAAATATTAATATATAGAACTAAGACAAAGCCCATTTGTT 120
Db 698 GAACGATCACAAACAAAATATTAAGTATATAGAACAAACAAAGCCCATCTGTT 757
QY 121 TTATATCTGAGAGATGCTCCAGCTACCCAAAATTAATAGAGAGTCCAGAGAA 180
Db 758 CTACATACCTGGAGAGTGTCCACAGAGCCAAAAGTTTGTGATGATCTCAGAAAA 817
QY 181 AATGAACGCTTATTTGAAGGTAGACGCAATGCAATTTGCAACAAATTAATTAATGGA 240
Db 818 ACTTAATGCAATATTTGAAGGTAGAGCGGCTTGAAGTTTGTGAGAGCTTAAACAAATGGA 877
QY 241 GGCTAGGCTAGAGACCAATCAATGAAAGAAAGAGCATCAG-----GTGTCG 291
Db 878 AGCTAGGCTAGAGCGCTAGCTCAGTGAAGAGAGACCAAGAGAGAGAGAGAGAGAGCA 937
QY 292 TAATGAAGACAGAGCGGAGCAACAGAGAGAGGTAGGCTGCTCAGCGAGAGAGAGT 351
Db 938 GAATGAAGACATACGAGAGATCGAGAGAGATGTAAGATGTCTCGCATTAATAGAGA--- 994
QY 352 GAGAGACAGAGTAATGACCAATGATGATGATGATGATGATGATGATGATGATGATGAT 411
Db 995 GGTGAGAGCGGTAAACCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1054
QY 412 AAAGAAATAGCATGTTTCATAGTATGATGATGATGATGATGATGATGATGATGATGAT 471
Db 1055 AGAGGTTA---CTATGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1111
QY 472 ACAGAAATGAGGTAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 531
Db 1112 AGATCTAGAGATTTAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1171
QY 532 GGATAGTACGCTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 591
Db 1172 AGAAGATCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1231
QY 592 AATTGCTGACAGAGAGATTAAGAACTAATGATGATGATGATGATGATGATGATGATGAT 651
Db 1232 CAAG 1291
QY 652 TGAAGCTAGCAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 711
Db 1292 TGAGATTAGCAATCAGAGATTAAGACTT-----GATGAGAGAT 1332
QY 712 TAAATCCCTTCTCTCGGAGAGAGAGATGATGATGATGATGATGATGATGATGATGAT 771
Db 1333 TAAA-----GGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1381
QY 772 GGAAG 831
Db 1382 ATTGAAGCTGAG 1441
QY 832 GTCTCAGCCCCAGCTTCAATCCAGTCCCAACCAAGTACTCCAGTCCCAAGCTCC 891
Db 1442 AGAGAGACTACAG 1501
QY 892 CTCTCAGCCTGAGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 951
Db 1502 GACTCAGCCAGA-----ACACCACTTAATGTTGGA 1531
QY 952 GCAAGGAGATTTACTACCTGAGAGAGAGAGATTTCTGATAGAGTCTGTAAGACTCAG 1011
Db 1532 TGAAGAGCCTTTC---CCTGAAAAAGAGAGAGATGCTGAGTGAATG---TGACTGA 1582

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 16, 2002, 12:50:13 : Search time 65 seconds
(without alignments)
6802.159 Million cell updates/sec

Title: US-09-822-849A-1
Perfect score: 1800
Sequence: 1 acagacagaactgcgcgtt.....ataaaaaaagaltacatcc 1800

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues
Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1730.4	96.1	2369	4	US-08-910-925-2	Sequence 2, Appli
2	125.6	7.0	7218	1	US-08-232-463-14	Sequence 14, Appli
3	121.2	6.7	3489	2	US-08-728-323A-1	Sequence 1, Appli
4	121.2	6.7	32207	2	US-08-770-379-20	Sequence 20, Appli
5	121.2	6.7	32207	4	US-08-757-669A-20	Sequence 20, Appli
6	121.2	6.7	32207	4	US-09-230-371A-20	Sequence 20, Appli
7	102.6	5.7	397	3	US-09-253-691-3	Sequence 3, Appli
8	99.2	5.5	16442	3	US-08-781-891-208	Sequence 208, App
9	97.2	5.4	3337	1	US-08-072-610-1	Sequence 1, Appli
10	97.2	5.4	3337	2	US-08-719-822B-1	Sequence 1, Appli
11	97.2	5.4	3337	4	US-09-092-458-1	Sequence 1, Appli
12	89.8	5.0	234	1	US-08-469-802B-3	Sequence 3, Appli
13	89.8	5.0	234	2	US-08-267-803B-3	Sequence 3, Appli
14	89.8	5.0	543	6	5273901-6	Patent No. 5273901
15	86.6	4.8	203	4	US-09-043-303-7	Sequence 7, Appli
16	81.8	4.5	533	6	5482709-5	Patent No. 5482709
17	81.2	4.5	5511	3	US-08-928-361B-2	Sequence 2, Appli
18	81.2	4.5	7334	3	US-08-928-361B-1	Sequence 1, Appli
19	81	4.5	195	1	US-08-469-802B-2	Sequence 2, Appli
20	81	4.5	195	2	US-08-267-803B-2	Sequence 2, Appli
21	80.6	4.5	51259	3	US-08-781-891-209	Sequence 209, App
22	79.8	4.4	165	4	US-09-043-303-17	Sequence 17, Appli
23	79.4	4.4	2580	3	US-09-050-863-2	Sequence 2, Appli
24	79.4	4.4	2580	4	US-09-359-081-2	Sequence 2, Appli
25	79.4	4.4	5452	2	US-09-130-114-1	Sequence 1, Appli
26	79.4	4.4	9600	4	US-08-910-647-1	Sequence 1, Appli
27	79.4	4.4	10596	1	US-07-884-811-15	Sequence 15, Appli

28	79.4	4.4	10596	1	US-07-885-971-15	Sequence 15, Appli
29	79.4	4.4	10596	1	US-08-087-783A-15	Sequence 15, Appli
30	79.4	4.4	10596	1	US-08-194-088B-15	Sequence 15, Appli
31	79.4	4.4	10596	2	US-08-194-087-15	Sequence 15, Appli
32	79.4	4.4	10596	5	PCT-US93-04648-15	Sequence 15, Appli
33	79	4.4	477	4	US-09-135-994-1	Sequence 1, Appli
34	75.2	4.2	154	1	US-08-469-802B-6	Sequence 6, Appli
35	75.2	4.2	154	2	US-08-267-803B-6	Sequence 6, Appli
36	74.2	4.1	3376	3	US-08-320-559-29	Sequence 29, Appli
37	74.2	4.1	3376	3	US-08-545-860D-29	Sequence 29, Appli
38	74.2	4.1	3376	5	PCT-US94-04496-29	Sequence 29, Appli
39	74	4.1	171	1	US-08-469-802B-5	Sequence 5, Appli
40	74	4.1	171	2	US-08-267-803B-5	Sequence 5, Appli
41	72	4.0	168	1	US-08-469-802B-4	Sequence 4, Appli
42	72	4.0	168	2	US-08-267-803B-4	Sequence 4, Appli
43	71.8	4.0	1276	4	US-09-177-325-2	Sequence 2, Appli
44	71.8	4.0	1276	4	US-09-411-812A-2	Sequence 2, Appli
45	71.8	4.0	1276	4	US-09-590-113-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-910-925-2
Sequence 2, Application US/08910925
Patent No. 6162601
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Shah, Purvi
TITLE OF INVENTION: HUMAN PININ SPLICE VARIANT
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910,925
FILING DATE: Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0365 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2369 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: FIBRN0T01
CLONE: 53219
US-08-910-925-2
Query Match 96.1%; Score 1730.4; DB 4; Length 2369;
Best Local Similarity 99.7%; Pred. No. 0;

		Matches 1734;	Conservative 0;	Mismatches 6;	Indels 0;	Gaps 0;
OY	1	ACAGACAGAACTGGCGCTTTTGGAAACAGAAAGTTGAGCTTGGCAGCTGCAGAGAGAAATG	60			
Db	630	ACAGACAGAACTGGCGCTTTTGGAAACAGAAAGTTGAGCTTGGCAGCTGCAGAGAGAAATG	689			
OY	61	GAATGAACATTAATGCCAAATTAATTAATATATAGAACTAAGACAAAGCCCATTTGTT	120			
Db	690	GAATGAACATTAATGCCAAATTAATTAATATATAGAACTAAGACAAAGCCCATTTGTT	749			
OY	121	TTATATTCCTGGAAGAAATGTGTCCAGCTACCCAAAATAATAGAGAGTCCAGAGAAA	180			
Db	750	TTATATTCCTGGAAGAAATGTGTCCAGCTACCCAAAATAATAGAGAGTCCAGAGAAA	809			
OY	181	AATGAACGCTTATTTAGAGTAGACGATCGAATTTGCAGAACAAATTAATAATGA	240			
Db	810	AATGAACGCTTATTTAGAGTAGACGATCGAATTTGCAGAACAAATTAATAATGA	869			
OY	241	GGCTAGGCTTAGAAGACAAATCAATGAAGGAAAAAGACATCAGGTGTCGTAATGAAGA	300			
Db	870	GGCTAGGCTTAGAAGACAAATCAATGAAGGAAAAAGACATCAGGTGTCGTAATGAAGA	929			
OY	301	ACAGAAAGCGGAAACAGAGAGGTTAGGTGCTCAGCGAGAGAGAAAGTTGAGAGAGAC	360			
Db	930	ACAGAAAGCGGAAACAGAGAGGTTAGGTGCTCAGCGAGAGAGAAAGTTGAGAGAGAC	989			
OY	361	AGGTAATCAGCACAATGATGTAGAAATAGAGAAAGCAGAGAGAGAAAGAAAT	420			
Db	990	AGGTAATCAGCACAATGATGTAGAAATAGAGAAAGCAGAGAGAGAAAGAAAT	1049			
OY	421	AGGATTTGTCATAGTAGAGAGAAAGAAACAGAGAGAGAGAAACAAACAGAAAT	480			
Db	1050	AGGATTTGTCATAGTAGAGAGAAAGAAACAGAGAGAGAGAAACAAACAGAAAT	1109			
OY	481	GGAGTTAAGATGAGAGAGAAACTGAGTTAAGGAAAGTGAAGACAGCAGATAGTCA	540			
Db	1110	GGAGTTAAGATGAGAGAGAAACTGAGTTAAGGAAAGTGAAGACAGCAGATAGTCA	1169			
OY	541	GGCTGAAGAAGTTATGATGTGCTAGAGATGTTGAGAAATGTCAAAACATGTAATGCTGA	600			
Db	1170	GGCTGAAGAAGTTATGATGTGCTAGAGATGTTGAGAAATGTCAAAACATGTAATGCTGA	1229			
OY	601	CCAGAGCTAATGGAACCTAATCCAGTTGAAGTGTAGAACCTTCAGAAATGAACCTAG	660			
Db	1230	CCAGAGCTAATGGAACCTAATCCAGTTGAAGTGTAGAACCTTCAGAAATGAACCTAG	1289			
OY	661	CAAGAATTTGGAACCAAGAAATTTGAATTTGAGCCAGATTAAGAATGTAAATCCCT	720			
Db	1290	CAAGAATTTGGAACCAAGAAATTTGAATTTGAGCCAGATTAAGAATGTAAATCCCT	1349			
OY	721	TTCTCTGGGAAAGAGAAATGTCTAGTCTTTAGACATGGAAGAGAGTCTGAGGAAAGA	780			
Db	1350	TTCTCTGGGAAAGAGAAATGTCTAGTCTTTAGACATGGAAGAGAGTCTGAGGAAAGA	1409			
OY	781	AGAAAAAGAAATCTGAGCCCAACCTGAGCCTGTGGCTCAACCTCAGCCTCAGTCCACC	840			
Db	1410	AGAAAAAGAAATCTGAGCCCAACCTGAGCCTGTGGCTCAACCTCAGCCTCAGTCCACC	1469			
OY	841	CCAGCTCAGCTTCAATCCCAAGTCCCAACAGTACTCCAGTCCCAAGCCTCCTCAGCC	900			
Db	1470	CCAGCTCAGCTTCAATCCCAAGTCCCAACAGTACTCCAGTCCCAAGCCTCCTCAGCC	1529			
OY	901	TGAGCAATTTGTCATAGCTGTTTACAGCCCAACACCCCAAGTTACTCAGGAGCAAGGCA	960			
Db	1530	TGAGCAATTTGTCATAGCTGTTTACAGCCCAACACCCCAAGTTACTCAGGAGCAAGGCA	1589			
OY	961	TTTACTACTGAGAGGAGGATTTTCTGTAGAGTCTGTAAACTCACTGAGGTACCACT	1020			
Db	1590	TTTACTACTGAGAGGAGGATTTTCTGTAGAGTCTGTAAACTCACTGAGGTACCACT	1649			
OY	1021	AGAGCCAGTCTTGACAGTACATCCAGAGAGCAAGAGCAAAACCAAACTAGAGCAGAG	1080			
Db	1650	AGAGCCAGTCTTGACAGTACATCCAGAGAGCAAGAGCAAAACCAAACTAGAGCAGAG	1709			

OY	1081	TAGAGTCGACCTAGAAATAAACCAAGCAGTAGAAGTCGAAGCAGTAGCAGTAGCAG	1140
Db	1710	TAGAGTCGACCTAGAAATAAACCAAGCAGTAGAAGTCGAAGCAGTAGCAGTAGCAG	1769
OY	1141	TTCTAGTAGAGTTCACACAGTAGCAGCAGTGAAGTAGTTCACAGCAGTGAAGTAGTAG	1200
Db	1770	TTCTAGTAGAGTTCACACAGTAGCAGCAGTGAAGTAGTTCACAGCAGTGAAGTAGTAG	1829
OY	1201	CAGTCGAGTAGTTCAGTAGCAGTCCAGTACAAAGTGGCAGCAGCAGAGATAGTAG	1260
Db	1830	CAGTCGAGTAGTTCAGTAGCAGTCCAGTACAAAGTGGCAGCAGCAGAGATAGTAG	1889
OY	1261	CAGTAGCAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAG	1320
Db	1890	CAGTAGCAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAG	1949
OY	1321	AGATAGAAAGCAGAGAGAGCGGTGATCCGAAGAGAGAGAGAGAGAGAGAGAGAG	1380
Db	1950	AGATAGAAAGCAGAGAGAGCGGTGATCCGAAGAGAGAGAGAGAGAGAGAGAGAG	2009
OY	1381	AAGTACCAATCTTCAAAAGGTGCTGTAGTAGAGATACAAAGAGATCAAGAGATAGAA	1440
Db	2010	AAGTACCAATCTTCAAAAGGTGCTGTAGTAGAGATACAAAGAGATCAAGAGATAGAA	2069
OY	1441	TTCCCGGTCCGACAGAAAGAGGTCTATATCAGAGAGTAGTCGATCAGGCAAAAGATCTTC	1500
Db	2070	TTCCCGGTCCGACAGAAAGAGGTCTATATCAGAGAGTAGTCGATCAGGCAAAAGATCTTC	2129
OY	1501	AAGAAGTGAAG	1560
Db	2130	AAGAAGTGAAG	2189
OY	1561	AGGCTTCTTAGCATTCTTTGACAGCAGAGATTTCTTGATTAATAAAGATTACCTTTCC	1620
Db	2190	AGGCTTCTTAGCATTCTTTGACAGCAGAGATTTCTTGATTAATAAAGATTACCTTTCC	2249
OY	1621	TTGTAAG	1680
Db	2250	TTGTAAG	2309
OY	1681	GACTGTTGTTTACAGACATCTTGTACTTTTTCATTAATTTTGAAGATTATTTATC	1740
Db	2310	GACTGTTGTTTACAGACATCTTGTACTTTTTCATTAATTTTGAAGATTATTTATC	2369

RESULT 2
 US-08-232-463-14/c
 ; Sequence 14, Application US/08232463
 ; Patent No. 5670367
 ; GENERAL INFORMATION:
 ; APPLICANT: DORNER, F.
 ; APPLICANT: SCHEIFLINGER, F.
 ; APPLICANT: FALKNER, F. G.
 ; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
 ; NUMBER OF SEQUENCES: 52
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 1800 diagonal Road, Suite 500
 ; CITY: Alexandria
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22313-0299
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/232,463
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:

Db 1459 CAGCAGAGCCACAGCAGAGAGCCCTGAGAGCCACAAACAGAGAGCCACAAACAG 1518
 QY 833 TCTAGCCCCAGCTTCAGCTTCATCCAGTCCCAACAGAGTACTCCAGTCCAGCTCCC 892
 Db 1519 CAGAGCCACAGCAGAGAGCCACAGCAGAGAGAGCCACAGCAGAGAGCCACAGCAG 1578
 QY 893 TCTAGCCTGAGAGATTGTCTATTAGCTGTTTACAGCCAAACCCCAAGTTACTCAGAG 952
 Db 1579 CAGAGCCACAGCAGAGAGAGCCACAGCAGAGAGAGCCACAGCAGAGAGCCACAGCAG 1638
 QY 953 CAAGGCATTACTACTGAGAGAGAGAGATTTCTGTAGAGTCTGTAAACTCACTGAG 1012
 Db 1639 CAGAGCCACAGCAGAGAGAGCCACAGCAGAGAGAGCCACAGCAGAGAGCCACAGCAG 1698
 QY 1013 GTACAGTAGAGCCAGCTTGTACAGTACATCCAGAG---AGCAAGAGCAAAACCAAAAC 1068
 Db 1699 CCGAGCCACAGCAGAGAGAGCCACAGCAGAGAGAGCCACAGCAGAGAGCCACAGCAG 1758
 QY 1069 TAGAGCAGAGTAGAGTGCAGCTAGAAATAAACAAGCAAGAGTAGAAGTGAAGCAG 1128
 Db 1759 CCGAGCCACAGCAGAGAGAGCCACAGCAGAGAGAGCCACAGCAGAGAGCCACAGCAG 1818
 QY 1129 TAGCAGTAGAGTGTCTAGTACAGTTCACACAGTACAGCAGAGTGAAGTGTCCAGCAG 1188
 Db 1819 CAGGATGAGCAGCAGAGAGTACAGCAGCAGAGAGTACAGCAGCAGAGATGACAGCAG 1878
 QY 1189 TCGAAGTAGTAGAGTCCAGTAGTCCAGTAGCAGCTCCAGTACAGTGAAGTGAAGCAGCAG 1248
 Db 1879 CAGGATGAGCAGCAGAGAGTACAGCAGCAGAGAGTACAGCAGCAGAGATGACAGCAG 1938
 QY 1249 CAGAGATAGTAGCAGTACAGTACAGTTCACACAGTACAGCAGAGTGAAGTGTCCAGCAG 1308
 Db 1939 CAGGATGAGCAGCAGAGAGTACAGCAGCAGAGAGTACAGCAGCAGAGATGACAGCAG 1998
 QY 1309 GGGACATATAGAGATAGAAAGCAGAGAGAGCGGTGATCCGAAGAGAGG 1360
 Db 1999 CAGGATGAGCAGCAGAGAGTACAGCAGCAGAGAGATGACAGCAGCAGAGATG 2050

RESULT 4

US-08-770-379-20/c

Sequence 20, Application US/08770379

Patent No. 5849564

GENERAL INFORMATION:

APPLICANT: Chang, Yuan

APPLICANT: Bohenzky, Roy A.

APPLICANT: Russo, James J.

APPLICANT: Edelman, Isidore S.

APPLICANT: Moore, Patrick S.

TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESS: Cooper & Dunham LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/770,379

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 52342

TELECOMMUNICATION INFORMATION:

TELECOMMUNICATION INFORMATION:

: TELEPHONE: (212) 278-0400
 : TELEFAX: (212) 391-0525
 : INFORMATION FOR SEQ ID NO: 20:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 32207 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: double
 : TOPOLOGY: linear
 : MOLECULE TYPE: DNA (genomic)
 : US-08-770-379-20

Query Match 6.7%; Score 121.2; DB 2; Length 32207;
 Best Local Similarity 45.9%; Pred. No. 2.3e-18;
 Matches 492; Conservative 0; Mismatches 573; Indels 7; Gaps 2;

QY 293 AATGAAGACAGCAAGCGGACACAGAGAGGTAAGTGGCTCAGCCAGAGAGAGAGTTG 352
 Db 21015 AATGAAGATGATGAG 20956
 QY 353 GAGAGACAGGTAATCAGCACAATGATGTAAGAAATAGAGAGAGAGAGAGAGAGAA 412
 Db 20955 GATGACGAGAGAGATGACGAGAGAGATGACGAGAGAGAGAGAGAGAGATGACGAGAG 20896
 QY 413 AAGAAATAGCGATTTCTCATAGTATGATGACAGAGAAAGAGAGAGAGAGAGAGAG 472
 Db 20895 GATGACGAGAGAGATGACGAGAGAGATGACGAGAGAGAGAGAGAGATGACGAGAGAG 20836
 QY 473 CAGGAATGAGGTTAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 532
 Db 20835 GAGGACGAT 20776
 QY 533 GATAGTCAAGCTGAG 592
 Db 20775 GATGAGGACATGAG 20716
 QY 593 ATGCTGACAGCAGAGAGTATGAGAACTAATCGAGTGAAGTGTAGAACTTCAGAAAT 652
 Db 20715 GACGGGGGC---GATGAAACAAACGTTGAGCATCCAAAGTTCAACACAGCAGCAGAG 20659
 QY 653 GAAGCTAGCAAGAAATTGGAACCAAGAAATGGAATTTGAAATTCAGCCAGATTAAGATGT 712
 Db 20658 CCACACAGCAGAGAGCCACAGCAGCAGAGAGCCACAGCAGCAGAGAGAGAGAGAGCC 20599
 QY 713 AATCCCTTTCTCTGAGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 772
 Db 20598 CAACAGCAGAGAGCCACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20539
 QY 773 GAAAAAGAGAAAAAGAAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 832
 Db 20538 CAGCAGAGAGCCACAGCAG 20479
 QY 833 TCTAGCCCCAGCTTCAGCTTCATCCAGTCCCAACAGTACTCCAGTCCAGCTCCC 892
 Db 20478 CAGGAGCCACAGCAGAGAGAGCCACAGCAGAGAGAGCCACAGCAGAGAGAGCCACAGCAG 20419
 QY 893 TCTAGCCTGAGAGATTGTCTATTAGCTGTTTACAGCCAAACCCCAAGTTACTCAGAGAG 952
 Db 20418 CAGGAGCCACAGCAGAGAGAGCCACAGCAGAGAGAGCCACAGCAGAGAGAGCCACAGCAG 20359
 QY 953 CAAGGCATTACTACTGAGAGAGAGAGATTTCTGTAGAGTCTGTAAACTCACTGAG 1012
 Db 20358 CAGGAGCCACAGCAGAGAGAGCCACAGCAGAGAGAGCCACAGCAGAGAGCCACAGCAG 20299
 QY 1013 GTACAGTAGAGCCAGCTTGTACAGTACATCCAGAG---AGCAAGAGCAAAACCAAAAC 1068
 Db 20298 CCGGAGCCACAGCAGAGAGAGCCACAGCAGAGAGAGCCACAGCAGAGAGCCACAGCAG 20239
 QY 1069 TAGGAGCAGAGTAGAGTGCAGCTAGAAATAAACAAGCAAGAGTAGAAGTGAAGCAGAG 1128
 Db 20238 CCGGAGCCACAGCAGAGAGAGCCACAGCAGAGAGAGTACAGCAGCAGAGATGACAGCAG 20179
 QY 1129 TAGCAGTAGAGTGTCTAGTACAGTTCACACAGTACAGCAGAGTGAAGTGTCCAGCAG 1188

Db 20178 CAGGATGAGCAGCAGCAGATGAGCAGCAGCAGATGAGCAGCAGCAGATGAGCAGCAG 20119
QY 1189 TGAAGTAGTAGCAGTCGACAGTAGTTCAGTAGCAGCTCCAGTACAGTGGCAGCAGCAG 1248
Db 20118 CAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGCAGATGAGCAGCAGCAGCAGCAG 20059
QY 1249 CAGAGTAGTAGCAGTAGCAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAG 1308
Db 20058 CAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGCAGATGAGCAGCAGCAGCAGCAGCAG 19999
QY 1309 GGGACATATAGATAGAGAGCAGCAGAGAGCGCTGATCGGAAGAGAGG 1360
Db 19998 CAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGCAGATGAGCAGCAGCAGCAGATG 19947

RESULT 5

US-08-757-669A-20/c
; Sequence 20, Application US/08757669A
Patent No. 6183751

GENERAL INFORMATION:

APPLICANT: Chang, Yuan
APPLICANT: Bohenzky, Roy A.
APPLICANT: Russo, James J.
APPLICANT: Edelman, Isidore S.
APPLICANT: Moore, Patrick S.
TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
TITLE OF INVENTION: SEQUENCES AND USES THEREOF
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757,669A
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 45185-F
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 32207 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-757-669A-20

Query Match 6.7%; Score 121.2; DB 4; Length 32207;
Best Local Similarity 45.9%; Pred. No. 2.3e-18;
Matches 492; Conservative 0; Mismatches 573; Indels 7; Gaps 2;

QY 293 AATGAAGACAGAGCGGGAACAAGAGAGGTTAGGTGGCTCAGCAGAGAGAGAGTGTG 352
Db 21015 AATGAAGATGATGAGGAGAGCAGAGACAGATGAGAGAGAGAGAGAGATGAGCAGAG 20956
QY 353 GAGGAGACAGGTATATCAGCAATGATGTAGAAATAGAGAGAGAGAGAGAGAGAGAA 412
Db 20955 GATGACGAGAGATGAGAGAGAGATGACGAGAGAGATGACGAGAGAGATGACGAGAG 20896
QY 413 AAGGAATAGCATTTCTCATAGTATGATGACAGAGAAAGAACAGAGAGAGAGAGAGAA 472

Db 20895 GATGACGAGAGAGATGACGAGAGAGAGATGACGAGAGATGACGAGAGAGATGACGAGAG 20836
QY 473 CAGGAATGAGAGTTAAGATGAGAGAGAGAGAACTGAGTTAAGGAAAGTGAAGACAGCAG 532
Db 20835 GAGGACGATGACGAT 20776
QY 533 GATAGTACGCTGAGAGAGTTATGATGTGCTAGAGATGTTGAGAAATGTAACATGTA 592
Db 20775 GATGAGACCAATGAGAGAGAGAGATGAGCAGAGAGAGAGAGAGAGAGAGAGAGAG 20716
QY 593 ATTGCTGACCAGAGAGTTAATGGAACCTAATCGAGTTGAAGTGAAGTGAAGTGAAGTGA 652
Db 20715 GACGGGGC--GATGGAACAACAACCTTGAGCATCCAAGTTCCACAACAGCAGCAGAGAG 20659
QY 653 GAAGCTAGCAAGAAATTTGAACCCAGAAATGGAATTTGAATTTGAGCCAGATTAAGATGT 712
Db 20658 CCACAACAGCAGAGAGCCACAGCAGCAGAGAGCCACAGCAGAGAGCCCTGACAGAGAGCA 20599
QY 713 AATCCCTTCTCCTGGGAAAGAGATGTCAGTCTTTAGACATGGAAGAGAGTGTAG 772
Db 20598 CAACAGCAGAGAGCCACAGCAGCAGAGAGCCACAGCAGAGAGAGAGAGAGAGAGAGAG 20539
QY 773 GAAAAAGAGAAAAAGATCTGAGCCCAACCTGAGCCCTGTGGCTCAACCTCAGCCTCAG 832
Db 20538 CAGCAGAGAGCCACAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20479
QY 833 TCTCAGCCCGAGCTTCACTTCAATCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 892
Db 20478 CAGGAGCCACAGCAGCAGAGAGCCACAGCAGCAGAGAGCCACAGCAGAGAGAGAGAG 20419
QY 893 TCTCAGCCCTGAGATTTGATTTAGCTGTTTACAGCCCAACCCCAAGTTACTCAGGAG 952
Db 20418 CAGGAGCCACAGCAGCAGAGAGCCACAGCAGCAGAGAGCCACAGCAGAGAGAGAGAG 20359
QY 953 CAAGGCAATTACTACTGAGAGAGAGAGATTTCTGTAGAGTCTGTAAGTCACTGAG 1012
Db 20358 CAGGAGCCACAGCAGCAGAGAGCCACAGCAGCAGAGAGAGAGAGAGAGAGAGAGAG 20299
QY 1013 GTACCAGTAGAGCCAGTCTGACAGTACATCCAGAG--AGCAAGAGCAAAACCAAAAC 1068
Db 20298 CCGGAGCCACAGCAGCAGAGAGCCACAGCAGCAGAGAGCCACAGCAGAGAGAGAGAG 20239
QY 1069 TAGGAGCAGAGTAGAGTTCGAGCTAGAAATTAACAAGCAGAGTAGAGTAGAGTAGAG 1128
Db 20238 CCGGAGCCACAGCAGCAGAGAGCCACAGCAGCAGAGTAGAGCAGCAGAGTAGAGCAG 20179
QY 1129 TAGCAGTAGCAGTTCTAGTAGCAGTTCACAGCAGTTCAGCAGCAGTTCAGCAGCAG 1188
Db 20178 CAGGATGAGCAGCAGCAGAGTAGAGCAGCAGCAGCAGAGTAGAGCAGCAGAGTAGAG 20119
QY 1189 TGAAGTAGTAGCAGTTCGAGTAGTTCAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAG 1248
Db 20118 CAGGATGAGCAGCAGCAGAGTAGAGCAGCAGCAGCAGAGTAGAGCAGCAGAGTAGAG 20059
QY 1249 CAGAGTAGTAGCAGTAGCAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAG 1308
Db 20058 CAGGATGAGCAGCAGCAGAGTAGAGCAGCAGCAGCAGAGTAGAGCAGCAGAGTAGAG 19999
QY 1309 GGGACATATAGATAGAGAGAGCAGAGAGAGCGCTGATCGGAAGAGAGG 1360
Db 19998 CAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGCAGATGAGCAGCAGCAGCAGATG 19947

RESULT 6

US-09-230-371A-20/c
; Sequence 20, Application US/09230371A
Patent No. 6348586

GENERAL INFORMATION:

APPLICANT: Chang, Yuan
APPLICANT: Bohenzky, Roy A
APPLICANT: Russo, James J
APPLICANT: Edelman, Isidore S

APPLICANT: Moore, Patrick S
TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
FILE OF INVENTION: USES THEREOF
CURRENT APPLICATION NUMBER: US/09/230,371A
CURRENT FILING DATE: 1999-11-17
PRIOR APPLICATION NUMBER: PCT/US97/13346
PRIOR FILING DATE: 1997-07-22
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn ver. 2.0
SEQ ID NO 20
LENGTH: 32207
TYPE: DNA
ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-230-371A-20

Query Match 6.7%; Score 121.2; DB 4; Length 32207;
Best Local Similarity 45.9%; Pred. No. 2.3e-18;
Matches 492; Conservative 0; Mismatches 573; Indels 7; Gaps 2;

293 AATGAGACAGAGCGGAGACAGAGAGGTAAGCTGCTCAGCAGAGAGAGTTG 352
DB 21015 AATAGGATGATGAGAGAGAGAGAGAGAGATGAGAGAGAGAGAGAGAG 20956
OY 353 GAGAGACAGATATGACACATATGATGTAAGTAAGAGAGAGAGAGAGAA 412
DB 20955 GATGACGAGAGATGACGAGAGAGATGACGAGAGAGATGACGAGAGAG 20896
OY 413 AAGGAATAGCATGTTTCATAGTAGTGCAGAGAAAGACAGAGAGAGAGAA 472
DB 20895 GATGACGAGAGATGACGAGAGAGATGACGAGAGAGATGACGAGAGAG 20836
OY 473 CAGGAATGAGGTTAAGTAGAGAGAGAACTGAGTAAGGAAAGTGAACAG 532
DB 20835 GAGGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGAC 20776
OY 533 GATAGTCAAGCTGAGAGATGATGATGCTGAGAGATGTTGAGATGTAAC 592
DB 20775 GATGAGACATGAGAGAGAGAGAGATGACGAGAGAGAGAGAGAGAG 20716
OY 593 ATTGCTGACGAGAGATGATGAACTAATCGAGTTGAAAGTGTAGAACCT 652
DB 20715 GACGGGGGCG--GATGGAACAAACGTTGAGCATCCAAAGTTCACAA 20659
OY 653 GAAGCTAGCAAGATTTGGAACCAAGATTTGAAATTGAGCCAGATTAAGA 712
DB 20658 CCACACAG 20599
OY 713 AATCCCTTTCTCTGGAAGAGATGTCAGTCTTTAGACATGGAAGAGAGT 772
DB 20598 CAACAGCAG 20539
OY 773 GAAAAAGAGAAAAAGATCTGAGCCCCCACTGAGCCTGTGCTCAACCTC 832
DB 20538 CAGCAG 20479
OY 833 TCTCAGCCCTCAGCTTCAATCCAGTCCCAACAGTACTCCAGTCCAGCT 892
DB 20478 CAGGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20419
OY 893 TCTCAGCCTGAGATTTGTATTAGCTGTTTACAGCCAAACCCCAAGTT 952
DB 20418 CAGGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20359
OY 953 CAAGGCAATTACTACTGAGAGAGAGAGATTTCTGTAGAGTCTGTAAAC 1012
DB 20358 CAGGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20299
OY 1013 GTACAGTAGAGCAGCTTGACAGTAGATCAGAGAGAGAGAGAGAGAG 1068
DB 20298 CCGGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20239
OY 1069 TAGGAGCAGAGAGTAGAGTGCAGCTAGAAATATAAACAAGCAGAGAG 1128

DB 20238 CCGGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20179
OY 1129 TAGCAGTAGAGCTTCTAGTAGAGAGTTCACAGAGAGAGAGAGAGAG 1188
DB 20178 CAGGATGAGAGAGAGAGAGATGAGCAGAGAGAGAGAGAGAGAGAG 20119
OY 1189 TGAAGTAG 1248
DB 20118 CAGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20059
OY 1249 CAGAGATAG 1308
DB 20058 CAGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 19999
OY 1309 GGGACATTAATAGAGATGAAAGCAGAGAGAGAGAGAGAGAGAGAG 1360
DB 19998 CAGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 19947

RESULT 7
US-09-253-691-3
Sequence 3, Application US/09253691
Patent No. 6124100
GENERAL INFORMATION:
APPLICANT: Dong Kyu JIN
TITLE OF INVENTION: Diagnostic Method and Kit for Neuropsychiatric Diseases
FILE REFERENCE: 1942/36
CURRENT APPLICATION NUMBER: US/09/253,691
EARLIER FILING DATE: 1999-02-22
EARLIER APPLICATION NUMBER: KR 98-6,278
NUMBER OF SEQ ID NOS: 3
SOFTWARE: WordPerfect 6.1/Windows
SEQ ID NO 3
LENGTH: 397
TYPE: DNA
ORGANISM: human
US-09-253-691-3

Query Match 5.7%; Score 102.6; DB 3; Length 397;
Best Local Similarity 62.1%; Pred. No. 7.2e-15;
Matches 162; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

OY 1026 CAGTCTGACAGTATCATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1085
DB 100 CACTTTGAAATGTTTCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 159
OY 1086 GTCAGCTAGAAATATAAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1145
DB 160 GCAGCAG 219
OY 1146 GTAGCAGTTCAACAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1205
DB 220 GCAGCAG 279
OY 1206 GCAGTAGTTCCAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1265
DB 280 GCAGCAG 339
OY 1266 GCACTAGTAGTAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1286
DB 340 GCAGCAG 360

RESULT 8
US-08-781-891-208/c
Sequence 208, Application US/08781891
Patent No. 6090620
GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui
APPLICANT: Yu, Chang-En

APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.
APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
TITLE OF INVENTION: WERNER'S SYNDROME
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentia Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: No. 6090620lenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 208:
SEQUENCE CHARACTERISTICS:
LENGTH: 16442 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-781-891-208

Query Match 5.5%; Score 99.2; DB 3; Length 16442;
Best Local Similarity 56.0%; Pred. No. 1.9e-13;
Matches 188; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 1054 GACCAAAACCAAACTAGAGCAGAGTAGAGTGCAGCTAGAAATAAACAAGCAGAG 1113
DB 16441 GAGCAGAGAGAGAGAGCAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 16382
QY 1114 TAGAATCCAGCAGTAGCAGTAGCAGTTCTAGTAGCAGTTCAACCAAGTAGCAGTGG 1173
DB 16381 GAGCAGAGAGAGAGAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 16322
QY 1174 AAGTAGTCCAGCAGTAGTGAAGTAGTAGCAGTGCAGTAGTCCAGTAGCAGTCCAGTAC 1233
DB 16321 GAGCAG 16262
QY 1234 AAGTGCAGCAGCAGCAGAGAGATAGTAGCAGTAGCAGTAGTAGTAGTAGTAGAGTAGAG 1293
DB 16261 GAGCAG 16202
QY 1294 TCGAGTAGAGGGTCCGGACATATAGAGATAGAAAGCACAGAGAGCGGTGATCGGAA 1353
DB 16201 TAG 16142
QY 1354 GAGAGGATAGTCTCAGAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1389
DB 16141 AAG 16106

RESULT 9

US-08-072-610-1

Sequence 1, Application US/08072610

Patent No. 5532133

GENERAL INFORMATION:

APPLICANT: Barnwell, John

TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens,

TITLE OF INVENTION: Monoclonal Antibodies, and Diagnostic Assays
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby and Darby
STREET: 805 Third Ave.
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022-7513

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentia Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/072,610
FILING DATE: 19930602
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gogoris, Adda
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 5986/07686
TELEPHONE: (212)527-7700
TELEFAX: (212)753-6237
TELEX: 236687

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 3337 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Plasmodium vivax
IMMEDIATE SOURCE:
CLONE: pVMB3.3.1

US-08-072-610-1

Query Match 5.4%; Score 97.2; DB 1; Length 3337;
Best Local Similarity 51.7%; Pred. No. 2.9e-13;
Matches 247; Conservative 0; Mismatches 228; Indels 3; Gaps 1;

QY 1033 GACAGTACATCCAGAGAGCAGAGCAAAACCAAACTAGAGAGCAGAGTAGAGGTCCAGC 1092
DB 1957 GAGAGCCAAACAG 2016
QY 1093 TAGAATTAACCAAGCAGAGAGTAGAGAGTGCAGAGTAGCAGTAGCAGTTCTAGTAGCAG 1152
DB 2017 TCGAATTAG 2076
QY 1153 TTCAACAGTAGCAGAGAGAGAGTAGTAGTCCAGCAGTGCAGTAGTAGTAGTAGTAGTAG 1212
DB 2077 CAG 2136
QY 1213 TTCAGTAGCAGCTCCAGTAGCAAGT--GGCAGCAGCAGCAGAGAGATAGTAGCAGTAGCAC 1269
DB 2137 CAGAGTAGAGAGAGAGAGAGAGTAGCAGAGAGAGTAGAGAGAGTAGAGAGAGTAGAG 2196
QY 1270 TAGTAGTAGTAGTAG 1329
DB 2197 AAGAGTGAAGAGAGTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2256
QY 1330 GCACAGAGAGAGCGGTGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1389
DB 2257 CAG 2316
QY 1390 ATCTTCAAAAGGTGCTGAGTAGAGATACAAAGAGATCAAGAGATTAAGAAATCCCGGTC 1449
DB 2317 AGGTACCAAGCGGTAGTAGAGTAGAGAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2376

Db	1957	GAGAGCCACAGAAAGGAGAGTGC	CAGAGAGAAATTAGAGGC	AACTCCAGAGCAGATT	2016
Qy	1093	TAGAAATAAAAACAAGCAAGAGT	AGAGTTCGAACGAGTAGCAGT	AGCTTCTAGTAGCAG	1152
Db	2017	TCGAATTAGAAACAACCAACAGAG	AAGAGTAGAAGAAACCGTAGAGG	CGCAAGAAACTG	2076
Qy	1153	TTCAACCAAGTAGCAGCAGTGGAG	TAGTTCCAGCAGTGGAAAGTAGT	AGCAGTCGCACTG	1212
Db	2077	CAGAAAGAGAAAGTGGAAAGAGT	ACTCTGCAGAAAGTAGAAGAGT	GGAAGAGTACTTG	2136
Qy	1213	TTCCAGTAGCAGCTCCAGTACAAGT	--GGCAGCAGCAGCAGAGATAGT	AGTACAGTAC	1269
Db	2137	CAGAAAGTAGAAGAGTGGAAAGAGT	TACCAGAAAGAGTAGAAGAGT	TACCCGCAAGTAG	2196
Qy	1270	TAGTAGTAGTAGTGCAGAGTAGAAG	TCCGAGTAGCGGTCGGGACAT	AATAGAGATAGAAA	1329
Db	2197	AAGAAAGTGAAGAGGTACCAGAAAG	AGTGGAAAGGTACCAGAAAG	AGTGGAAAGGTAC	2256
Qy	1330	GCACAGAAAGCGCTGGATCGGAAGA	GAAAGGATCTTCAGGACTAGAAAG	AGTGCACAA	1389
	2257	CAGAAAGAGTGGAAAGGTACCAGAA	AGAGTGGAAAGTAGAAGAGTAG		2316
Qy	1390	ATCTTCAAAAGTGCTGCTAGTAGAG	ATACAAAAAGGATCAAAAGAT	TAAGAATTCCCGTC	1449
Db	2317	AGGTACCAAGCGTAGTAGAAGTAGA	AGTACCAGCGGTAGTAGAAGAAG	AGTGCAGAAAG	2376
Qy	1450	CGACAGAAAGAGGTATATATCAGAG	AGTAGTGCATCAGCCAAAGATCTT	CAAGAAAT	1507
Db	2377	AAGTAGAAGAAAGAAAGAGCAAGAA	CCAGTAGAGCAAGAGATGTATT	TACAATT	2434

RESULT 12
 US-08-469-802B-3
 : Sequence 3, Application US/08469802B
 : Patent No. 5741645
 :
 : GENERAL INFORMATION:
 : APPLICANT: Orr, Harry T.
 : APPLICANT: Ranum, Laura P.W.
 : APPLICANT: Chung, Ming-Yi
 : APPLICANT: Zoghbi, Huda Y.
 : TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
 : Patent No. 5741645
 :
 : TITLE OF INVENTION: Type 1 and Method for Diagnosis
 : NUMBER OF SEQUENCES: 47
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Muelling, Raasch, Gebhardt & Schwappach, P.A.
 : STREET: 119 No. 5741645th Fourth Street, Suite 203
 : CITY: Minneapolis
 : STATE: MN
 : COUNTRY: USA
 : ZIP: 55401
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 :
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 :
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/469,802B
 : FILING DATE: 06-JUN-1995
 : CLASSIFICATION: 435
 :
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Muelling, Ann M.
 : REGISTRATION NUMBER: 33,977
 : REFERENCE/DOCKET NUMBER: 110.00030101
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 612-305-1217
 : TELEFAX: 612-305-1225
 :
 : INFORMATION FOR SEQ ID NO: 3:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 234 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear

```

; MOLECULE TYPE: DNA
US-08-469-802B-3

Query Match          5.0%; Score 89.8; DB 1; Length 234;
Best Local Similarity 66.0%; Pred. No. 5.1e-12;
Matches 130; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 1106 AGCAAGAGTAGAAGTCGAAGCAGTAGCAGTAGTCTAGTAGCAGTTCACACAGTAGC 1165
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 61

QY 1166 AGCAGTGGAAAGTAGTTCACAGCAGTGAAGTAGTAGCAGTTCGACAGTAGTTCACAGTAGCAGC 1225
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 62 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 121

QY 1226 TCCAGTAGCAAGTGGCAGCAGCAGCAGCAGAGATAGTAGCAGTAGCAGCACTAGTAGTAGTGAG 1285
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 122 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 181

QY 1286 AGTAGAAGTCGAGTAG 1302
      || | | | | | | | |
Db 182 AGCAGCAGCAGCAGCAGCAG 198

```

```

RESULT 13
US-08-267-803B-3
; Sequence 3, Application US/08267803B
; Patent No. 5834183
; GENERAL INFORMATION:
; APPLICANT: Orr, Harry T.
; APPLICANT: Ranum, Laura P.W.
; APPLICANT: Chung, Ming-yl
; APPLICANT: Zoghbi, Huda Y.
; TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
; Patent No. 5834183
; TITLE OF INVENTION: Type 1 and Method for Diagnosis
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Muetting, Raasch, Gebhardt & Schwappach, P.A.
; STREET: P.O. Box 581415
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55458-1415
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/267,803B
; FILING DATE: 28-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McCormack, Myra H.
; REGISTRATION NUMBER: 36,602
; REFERENCE/DOCKET NUMBER: 110.00030120
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-305-1217
; TELEFAX: 612-305-1228
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 234 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
;
US-08-267-803B-3

```


[illegible]

RESULT 14

: Patent No. 5273901
 :
 : APPLICANT: JACOBSON, JAMES W.: STRAUSBERG, ROBERT L.: WILSON,
 : USAN D.; POPE, SHARON H.; STRAUSBERG, SUSAN L.; RUFF, MICHAEL D.;
 : AUGUSTINE, PATRICIA C.; DANFORTH, HARRY D.
 : TITLE OF INVENTION: GENETICALLY ENGINEERED COCCIDIOSIS
 : SPOROZOITE 21.5 KB ANTIGEN, AC-6B
 : NUMBER OF SEQUENCES: 11
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/07/581,693
 : FILING DATE: 12-SEP-1990
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 215,162
 : FILING DATE: 05-JUL-1988
 : APPLICATION NUMBER: 746,520
 : FILING DATE: 19-JUN-1985
 : APPLICATION NUMBER: 627,811
 : FILING DATE: 05-JUL-1984
 : SEQ ID NO: 6:
 : LENGTH: 543
 : 5273901-6

Query Match	5.0%;	Score 89.8;	DB 6;	length 543;
Best Local Similarity	56.7%;	Pred. No. 7.1e-12;		
Matches 166;	Conservative 0;	Mismatches 127;	Indels 0;	Gaps 0;

QY	1049	AGCAGAGCAAAAACCAAACCTAGGAGCAGAAGTAGAGCTGAGCTAGMAATAAACAAGC	1108
		11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11	
Db	21	AGAGGAAGACAAAAGAGAAAGAAAGAAAGAAAGAAAGAGAGAGAGAGAGAAAGAGAGA	80
		1109	
		AMGAGTAGAAGTCGAGCAGCAGTAGCAGTAGCAGTTCTAGTAGCAGTTCAACCACTAGCAGC	1168
		11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11	
	81	AAAAGAGAAAGAAAGAGCGGCAGCACCCTGCAGCAGCAACAGCAGCAGCACCCTGCAGCAGC	140
		1169	
QY		AGTGAAGTAGTTCAGCAGTAGTGAAGTAGTAGCAGTCCAGTAGTAGTCCAGTAGCAGCTCC	1228
		11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11	
Db	141	AMCAGCAGCAGCAACACAGCAGCTGCACACAGCAGCAGCAGCTACAGCAACACCAGCTGCAGCTGC	200
		1229	
QY		AGTACAAGTGGCAGCAGCAGCAGCAGAGATAGTAGCAGTAGCACTAGTAGTAGTAGTGAGAGT	1288
		11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11	
Db	201	AGCACACAGCTGCAGCAGCAGCAGCAGCTGCAGCAGCAGCAGCAGCAACAGAGAGCTGCAGCAGC	260
		QY	
	1289	AGAACTCGGAGTAGGGGCTCGGGGACATAATAGATAGAAAGCACAGAAAGAGG	1341
		11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11	
Db	261	ACCAGCTGCAGCAGCAGCCCGCAGCAGGAGCAGCAGCAGCAAAAGGCAGCGG	313

RESULT 15

US-09-043-303-7
: Sequence 7, Application US/09043303
: Patent No. 6251589
: GENERAL INFORMATION:
: APPLICANT: TSUJI, Shoji
: APPLICANT: SANPEI, Kazujiro
: TITLE OF INVENTION: Method, for Diagnosing Spinocerebellar Ataxia Type 2 and

```

; TITLE OF INVENTION: Primers Therefor
; FILE REFERENCE: 0760-0241P
; CURRENT APPLICATION NUMBER: US/09/043,303
; CURRENT FILING DATE: 1998-05-18
; EARLIER APPLICATION NUMBER: PCT/JP96/01999
; EARLIER FILING DATE: 1996-07-18
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 203
; TYPE: DNA
; ORGANISM: p-2093 plasmid
US-09-043-303-7

```

Query Match	4.88;	Score 86.6;	DB 4;	Length 203;
Best Local Similarity	65.08;	Pred. No. 2.6e-11;		
Matches 128;	Conservative 0;	Mismatches 69;	Indels 0;	Gaps 0;

OY	1103	ACAGCAGAAGTAGAAGTCCGACGACAGTAGCACAGTTCTAGTAGCAGTTCAACCAGT	1162
Dd	5	ACCAGCAACAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	64
OY	1163	AGCAGCAGTGGAAGTAGTTCACGACAGTGGAAAGTAGTAGCAGTCCGACAGTAGTTCACAGTAGC	1222
Dd	65	AGC	124
OY	1223	AGCTCCAGTACAAGTGGCAGCAGCAGCAGCAGAGATAGTAGCAGTACACTAGTAGTAGTAGT	1282
Dd	125	AGC	184
OY	1283	GAGAGTAGAAGTCGGAG	1299
Dd	185	ATCAGCGAAACTCTGGG	201

Search completed: October 16, 2002, 13:06:56
Job time : 158 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 16, 2002, 12:49:43 ; Search time 310 Seconds
(without alignments)
9969.181 Million cell updates/sec

Title: US-09-822-849A-1
Perfect score: 1800
Sequence: 1 acagacagaactgcgcgtt.....ataaaaaaagattacatcc 1800

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1730.4	96.1	2369	22	AAF24910 Nucleotide sequenc
2	607.4	33.7	616	21	AAA43765 Human secreted exp
3	393	21.8	1113	21	AAF15981 Human prostate can
4	299.4	16.6	325	20	AAV89627 EST clone COS59.
5	268.4	14.9	293	16	AAI20032 Human gene signatu
6	262.4	14.6	264	21	AAC02581 Human secreted pro
7	251.4	14.0	263	20	AAV86711 EST clone AW32. H
8	166.4	9.2	239	22	AA524061 Human ovarian PCR-
9	166.4	9.2	239	22	AAH82631 Human ovarian tumo

10	136.4	7.6	3399	17	AAT05868	Chicken leucocytoz
11	121.2	6.7	3489	21	AAA30290	Kaposi's sarcoma-a
12	121.2	6.7	3489	22	AAF82901	Nucleotide sequenc
13	121.2	6.7	32207	20	AAV73805	KSHV LUR DNA (nucl
14	121.2	6.7	137507	19	AAV19941	KSHV long unique c
15	118	6.6	1037	21	AAA59242	Exons E, C and A o
16	118	6.6	1159	21	AAA59240	Exons E, C and A o
17	118	6.6	1472	21	AAA59241	Exons D, C, B and
18	116	6.4	1686	16	AAO87587	Exons D, C, B and
19	103.6	5.8	5120	22	AAC84677	DNA encoding Leuco
20	102.6	5.7	397	20	AAV89891	DNA sequence of hu
21	99.2	5.5	16442	18	AAV83006	Spino cerebellar at
22	97.2	5.4	3337	17	AAV34620	Partial mouse WRN
23	97.2	5.4	3337	20	AAV15174	P. vivax ESP-1 blo
24	97.2	5.4	3337	22	AAV15174	DNA encoding a sec
25	92.4	5.1	1080	22	AAH76457	Plasmodium vivax E
26	92.4	5.1	575	22	ABA50472	Mouse ischaemic co
27	91.6	5.1	575	22	ABA68422	Human breast cell
28	91.6	5.1	575	22	ABA35413	Human foetal liver
29	91.6	5.1	575	22	AAK16793	Probe #13879 for g
30	91.6	5.1	575	22	AAK16793	Human brain expres
31	91.6	5.1	575	22	AAK42567	Human bone marrow
32	91.6	5.1	575	22	AAI48636	Probe #17322 used
33	91.6	5.1	1969	22	ABA45341	Probe #8947 used t
34	91.6	5.1	1969	22	ABA55830	Human breast cell
35	91.6	5.1	1969	22	ABA25506	Human foetal liver
36	91.6	5.1	1969	22	AAK04048	Probe #3972 for ge
37	91.6	5.1	1969	22	AAK29533	Human brain expres
38	91.6	5.1	1969	22	AAI14105	Human bone marrow
39	91.6	5.1	1969	22	AAI35486	Probe #4038 for ge
40	91.6	5.1	1969	22	AAI35486	Probe #4172 used t
41	91.6	5.1	1969	22	AAI03958	Probe #3949 used t
42	90.8	5.0	8201	21	AAA88864	Human dentin stalo
43	90.4	5.0	621	23	AAZ23895	DNA encoding novel
44	90.4	5.0	49999	20	AAZ23895	Murine LOBO homolo
45	89.8	5.0	234	16	AAQ84832	Spino cerebellar at

ALIGNMENTS

RESULT 1	
AAF24910	
ID	AAF24910 standard; cDNA; 2369 BP.
AC	AAF24910;
XX	
XX	
DT	20-APR-2001 (first entry)
XX	
DE	Nucleotide sequence of a human pinin splice variant (PNIN).
XX	
KW	Human; pinin splice variant; PNIN; vesicle trafficking; AIDS; emphysema;
KW	Addison's disease; neoplastic disorder; immunological disorder; gout;
KW	adult respiratory distress syndrome; allergy; anaemia; asthma; cancer;
KW	atherosclerosis; bronchitis; cholecystitis; Crohn's disease; scleroderma;
KW	ulcerative colitis; atopic dermatitis; dermatomyositis; osteoporosis;
KW	atrophic gastritis; diabetes mellitus; erythema nodosum; pancreatitis;
KW	glomerulonephritis; Graves' disease; hypereosinophilia; polymyositis;
KW	irritable bowel syndrome; lupus erythematosus; multiple sclerosis;
KW	myasthenia gravis; myocardial or pericardial inflammation; hemodialysis;
KW	osteoarthritis; rheumatoid arthritis; Sjogren's syndrome; trauma;
KW	Werner syndrome; autoimmune thyroiditis; infection; ss.
XX	
OS	Homo sapiens.
XX	
XX	
FH	Key
FT	CDS
FT	
FT	
FT	
XX	
PN	US6162601-A.
XX	
PD	19-DEC-2000.

Location/Qualifiers
25..2178
/*tag- a
/product- " pinin splice variant (PNIN)"

XX 08-AUG-1997; 97US-0910925.
XX 08-AUG-1997; 97US-0910925.
XX (IMCT-) INCYTE PHARM INC.
XX Bandman O, Shah P, Lal P;
XX WPI; 2001-090399/10.
DR P-PSDB; AAB31581.

Novel nucleotide sequence encoding human pinin splice variant, useful
for treating AIDS, Addison's disease, allergy, anaemia, asthma,
atherosclerosis and adult respiratory distress syndrome -

Claim 1; Fig 1A-G; 38pp; English.

The present sequence encodes a human pinin splice variant (PININ). PININ
polypeptides and polynucleotides are useful for treating, diagnosing and
preventing developmental, vesicle trafficking, neoplastic and
immunological disorders. Antagonists of PININ polypeptides are useful
for treating AIDS, Addison's disease, adult respiratory distress
syndrome, allergies, anaemia, asthma, atherosclerosis, bronchitis,
cholecystitis, Crohn's disease, ulcerative colitis, atopic dermatitis,
dermatomyositis, diabetes mellitus, emphysema, erythema nodosum, atrophic
gastritis, glomerulonephritis, gout, Graves' disease, hyperosinophilia,
irritable bowel syndrome, lupus erythematosus, multiple sclerosis,
myasthenia gravis, myocardial or pericardial inflammation,
osteoarthritis, osteoporosis, pancreatitis, polymyositis, rheumatoid
arthritis, scleroderma, Sjogren's syndrome, Werner syndrome, and
autoimmune thyroiditis, complications of cancer, hemodialysis, and
extracorporeal circulation, trauma, and viral, bacterial fungal,
parasitic, protozoal, and helminthic infections.

Sequence 2369 BP; 843 A; 406 C; 651 G; 469 T; 0 other;

Query Match 96.18; Score 1730.4; DB 22; Length 2369;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1734; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ACAGACAGAACTGCGCTTTTGAACAGAAAGTTGAGCTTCCGACGCTGCAAGAAAGATG 60
DB 630 ACAGACAGAACTGCGCTTTTGAACAGAAAGTTGAGCTTCCGACGCTGCAAGAAAGATG 689
QY 61 GAATGAACATATGCGCAAAATTAATATATATAAGAACTAAGCAAAAGCCCATTTGTT 120
DB 690 GAATGAACATATGCGCAAAATTAATATATATAAGAACTAAGCAAAAGCCCATTTGTT 749
QY 121 TTATATTCCTGGAAGATGTGTCACGCTACCCAAAACCTAATAGAAAGTCAAGAGAAA 180
DB 750 TTATATTCCTGGAAGATGTGTCACGCTACCCAAAACCTAATAGAAAGTCAAGAGAAA 809
QY 181 AATGAACGCTTTATTTGAAGTAGACGATCGAATTTGCAGAACAAATAAATGAATGA 240
DB 810 AATGAACGCTTTATTTGAAGTAGACGATCGAATTTGCAGAACAAATAAATGAATGA 869
QY 241 GGTAGGCTTAGAAGACAATCAATGAAGAAAAGAGCATCAGTGTGCGTAATGAAGA 300
DB 870 GGTAGGCTTAGAAGACAATCAATGAAGAAAAGAGCATCAGTGTGCGTAATGAAGA 929
QY 301 ACAGAAAGCGGGAACAGAGAGGTTAGGTGGCTCAGCGAGAGAAAGATTGGAGAGAC 360
DB 930 ACAGAAAGCGGGAACAGAGAGGTTAGGTGGCTCAGCGAGAGAAAGATTGGAGAGAC 989
QY 361 AGGTAATCAGACAATGATAGAAATAGAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAT 420
DB 990 AGGTAATCAGACAATGATAGAAATAGAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAT 1049
QY 421 AGCGATTGTCATAGTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 480
DB 1050 AGCGATTGTCATAGTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1109

QY 481 GGAGTTAAGATGAGAGAGAGAACTGAGGTAAGGAAAGTGAGAGACAGCAGATAGTCA 540
DB 1110 GGAGTTAAGATGAGAGAGAGAACTGAGGTAAGGAAAGTGAGAGACAGCAGATAGTCA 1169
QY 541 GCCTGAAGAGATTATGATGTGCTAGAGATGTTGAGAAATGTCAAACTATGTAATGCTGA 600
DB 1170 GCCTGAAGAGATTATGATGTGCTAGAGATGTTGAGAAATGTCAAACTATGTAATGCTGA 1229
QY 601 CCAGAGGTAATGAGAACTAATCGAGTTGAAGGTGTAGAACCTTCAGAAATGAAGCTAG 660
DB 1230 CCAGAGGTAATGAGAACTAATCGAGTTGAAGGTGTAGAACCTTCAGAAATGAAGCTAG 1289
QY 661 CAAGAATTGGAACCAAGAAATGGAATTTGAATTTGAGCCAGATTAAGAAATGTAATCCCT 720
DB 1290 CAAGAATTGGAACCAAGAAATGGAATTTGAATTTGAGCCAGATTAAGAAATGTAATCCCT 1349
QY 721 TTCTCCTGGAAAGAGATGTCAGTGTGCTTTAGACATGGAAGAGAGTCTGAGGAAAGA 780
DB 1350 TTCTCCTGGAAAGAGATGTCAGTGTGCTTTAGACATGGAAGAGAGTCTGAGGAAAGA 1409
QY 781 AGAAAAAGAAATCTGAGCCCAACCTGAGCCTGTGGCTCAACCTCAGCCTCAGTCAAGCC 840
DB 1410 AGAAAAAGAAATCTGAGCCCAACCTGAGCCTGTGGCTCAACCTCAGCCTCAGTCAAGCC 1469
QY 841 CCAGCTTCAGCTTCAATCCAGTCCCAACCAAGTACTCCAGTCCCAAGCCTCCTCAGCC 900
DB 1470 CCAGCTTCAGCTTCAATCCAGTCCCAACCAAGTACTCCAGTCCCAAGCCTCCTCAGCC 1529
QY 901 TGAGATTGTCATATAGTGTGTTTACAGCCACACACCCCAAGTTACTCAGAGCAAGGCA 960
DB 1530 TGAGATTGTCATATAGTGTGTTTACAGCCACACACCCCAAGTTACTCAGAGCAAGGCA 1589
QY 961 TTTACTACCTGAGAGAGAGATTTTCTGTAGAGTGTGTAATACTCAGTCAAGTACCAGT 1020
DB 1590 TTTACTACCTGAGAGAGAGATTTTCTGTAGAGTGTGTAATACTCAGTCAAGTACCAGT 1649
QY 1021 AGAGCCAGTCTTGACAGTACATCCAGAGAGCAAGCAAAACCAAACTAGAGCAGAGAG 1080
DB 1650 AGAGCCAGTCTTGACAGTACATCCAGAGAGCAAGCAAAACCAAACTAGAGCAGAGAG 1709
QY 1081 TAGAGGTGAGCTAGAAATTAACAAGCAAGAGAGTAGAAGTGAAGCAGTAGAGTAGCAG 1140
DB 1710 TAGAGGTGAGCTAGAAATTAACAAGCAAGAGAGTAGAAGTGAAGCAGTAGAGTAGCAG 1769
QY 1141 TTCTAGTAGCAGTTCAACCAAGTAGCAGCAGTGAAGTAGTTCCAGCAGTGAAGTAGTAG 1200
DB 1770 TTCTAGTAGCAGTTCAACCAAGTAGCAGCAGTGAAGTAGTTCCAGCAGTGAAGTAGTAG 1829
QY 1201 CAGTCGAGTAGTTCCAGTAGCAGCTCCAGTACCAAGTGGCAGCAGCAGCAGATAGTAG 1260
DB 1830 CAGTCGAGTAGTTCCAGTAGCAGCTCCAGTACCAAGTGGCAGCAGCAGCAGATAGTAG 1889
QY 1261 CAGTAGCACTAGTAGTAGTAGAGAGTAGAAGTGGAGTAGGGGTGGGACATTAATAG 1320
DB 1890 CAGTAGCACTAGTAGTAGTAGAGAGTAGAAGTGGAGTAGGGGTGGGACATTAATAG 1949
QY 1321 AGATAGAAAGCAGAGAGAGCGGTGATCGGAAGAGAGAGGATCTTCAGGACTAGAAAG 1380
DB 1950 AGATAGAAAGCAGAGAGAGCGGTGATCGGAAGAGAGAGGATCTTCAGGACTAGAAAG 2009
QY 1381 AAGTCACAAATCTTCAAAAGGTGTGTAGTAGAGATCAAAAAGGATCAAAAGATTAAGAA 1440
DB 2010 AAGTCACAAATCTTCAAAAGGTGTGTAGTAGAGATCAAAAAGGATCAAAAGATTAAGAA 2069
QY 1441 TTCCCGGTCCGACAGAAAGAGTCTATATCAGAGAGTAGTCATCAGCAAAAGATCTTC 1500
DB 2070 TTCCCGGTCCGACAGAAAGAGTCTATATCAGAGAGTAGTCATCAGCAAAAGATCTTC 2129
QY 1501 AAGAGTGAAGAGAGCCGGAATATCAGACAGGAAAGAGAGAGAGAGAGAGAGAGAGAG 1560
DB 2130 AAGAGTGAAGAGAGAGCCGGAATATCAGACAGGAAAGAGAGAGAGAGAGAGAGAGAG 2189
QY 1561 AGGCTTTCTTAGCCATTTCTTGACAGCAGAGATTTCTTGATTAAGAAAGATTAACCTTCC 1620

Db 2190 AGCTTTCTTAGCCATCTTTCAGCAGAGATTCTTGATTAAGGATTACCTTCC 2249
 QY 1621 TTGTAAGAGAGATGCTGCTTAAGATTCATGTTGTAAGAAATCTTTTGAAATACA 1680
 Db 2250 TTGTAAGAGAGATGCTGCTTAAGATTCATGTTGTAAGAAATCTTTTGAAATACA 2309
 QY 1681 GACTGTTGTTACAGACATCTTGTACTTTTGCAATATTTGTAAGAGTTATTAAC 1740
 Db 2310 GACTGTTGTTACAGACATCTTGTACTTTTGCAATATTTGTAAGAGTTATTAAC 2369

RESULT 2
 AAA43765
 ID AAA43765 standard; cDNA; 616 BP.

AC AAA43765;
 DT 21-APR-2000 (first entry)

Human secreted expressed sequence tag SEQ ID NO:340.

Human; mouse; chicken; rat; secreted expressed sequence tag; SEST; expressed sequence tag; EST; probe; chemotactic; proliferative; immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal; antiviral; antidiabetic; antiaesthetic; vulnery; antiparkinsonian; antitumor; osteopathic; neuroprotective; nootropic; antipsoriatic; cerebroprotective; anticonvulsant; antidepressant; gene therapy; vaccine; autoimmune disorder; multiple sclerosis; allergic condition; insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer; lymphoid cell deficiency; burn; osteoporosis; osteoarthritis; central nervous system disorder; Alzheimer's disease; stroke; Parkinson's disease; Huntington's disease; coagulation disorder; haemophilia; thrombosis; inflammatory disorder; Crohn's disease; tumour; infection; depression; psoriasis; ss.

Homo sapiens.

MO200021991-A1.

20-APR-2000.

15-OCT-1999; 99MO-US24206.

15-OCT-1998; 98US-0104436.

(GEMT) GENETICS INST INC.

Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C; Merberg D, Treacy M, Bowman MR;

WPI; 2000-317938/27.

Isolated polynucleotides, and encoded proteins, comprising secreted expressed sequence tags (SESTs), useful for treating various disorders such as autoimmune, infectious, and central nervous system disorders - Claim 1; Page 292; 803pp; English.

AAAA3426 to AAA45925 represent specifically claimed secreted expressed sequence tags (SESTs), isolated from human, mouse, chicken and rat tissue sources. The SESTs can have a range of activities depending on the tissues they were isolated from. The activities include: chemotactic; proliferative; immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal; antiviral; antidiabetic; antiaesthetic; vulnery; antiparkinsonian; antipsoriatic; cerebroprotective; nootropic; anticonvulsant; antidepressant. The SESTs can be used for gene therapy and in vaccines. The SESTs are useful as probes for the identification and isolation of full-length cDNAs and genomic DNA molecules which correspond to the SESTs. Proteins encoded by the SESTs

are useful in assays for determining biological activity and raising antibodies. They may be useful for treatment of autoimmune disorders (multiple sclerosis, insulin dependent diabetes), allergic conditions (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers, osteoporosis, osteoarthritis, central nervous system disorders (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's disease), tumours, bacterial, fungal or viral infections, depression and psoriasis. AAA45926 to AAA45931 represent linker variants which are given in the exemplification of the present invention.

Sequence 616 BP; 211 A; 126 C; 154 G; 125 T; 0 other;

Query Match 33.7%; Score 607.4; DB 21; Length 616;
 Best Local Similarity 99.8%; Pred. No. 1e-124;
 Matches 608; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 467 CAAAAACAGGAATGAGGTTAAGATGAGGAGGAAAGTAAGGAAAGTGAAG 526
 Db 2 CAAAAACAGGAATGAGGTTAAGATGAGGAGGAAAGTGAAGGAAAGTGAAG 61
 QY 527 CAGCAGATAGTACAGCCTGAGCAAGTTATGATGTGCTAGAGATGTTGAGATGTCAA 586
 Db 62 CAGCAGATAGTACAGCCTGAGCAAGTTATGATGTGCTAGAGATGTTGAGATGTCAA 121
 QY 587 CATGTAATGCTGACAGGAGGTAATGAAACTAATCGAGTTGAAAGTGTAGAACCTCA 646
 Db 122 CATGTAATGCTGACAGGAGGTAATGAAACTAATCGAGTTGAAAGTGTAGAACCTCA 181
 QY 647 GAAATATGAGCTAGCAAGAAATGGAACCAAGAAATGGAATTTGAAATTTGAGCCAGATAA 706
 Db 182 GAAATATGAGCTAGCAAGAAATGGAACCAAGAAATGGAATTTGAAATTTGAGCCAGATAA 241
 QY 707 GAATGTAATCCCTTTCTCTCTGGGAAAGAAATGTCAGTGTGTTAGACATGGAAGAGAG 766
 Db 242 GAATGTAATCCCTTTCTCTCTGGGAAAGAAATGTCAGTGTGTTAGACATGGAAGAGAG 301
 QY 767 TCTGAGGAAAAAGAAAGAAATGAGCCCAACCTGAGCCTGTGGCTCAACCTCAG 826
 Db 302 TCTGAGGAAAAAGAAAGAAATGAGCCCAACCTGAGCCTGTGGCTCAACCTCAG 361
 QY 827 CCTCAGTCTAGCCCCAGCTTCACTTCAATCCAGTCCCAACCACTACTCCAGTCCAG 886
 Db 362 CCTCAGTCTAGCCCCAGCTTCACTTCAATCCAGTCCCAACCACTACTCCAGTCCAG 421
 QY 887 CCTCCCTCAGCCTGAGGATTTGTATAGCTGTTTACAGCCCAACCACTACT 946
 Db 422 CCTCCCTCAGCCTGAGGATTTGTATAGCTGTTTACAGCCCAACCACTACT 481
 QY 947 CAGAGCAAGGCAATTTACTACCTGAGAGAGGATTTCTGTAGAGTCTGTAAACTC 1006
 Db 482 CAGAGCAAGGCAATTTACTACCTGAGAGAGGATTTCTGTAGAGTCTGTAAACTC 541
 QY 1007 ACTGAGTACAGTAGAGCCAGTCTTGACAGTACATCCAGAGAGCAAGCAAAACCAAA 1066
 Db 542 ACTGAGTACAGTAGAGCCAGTCTTGACAGTACATCCAGAGAGCAAGCAAAACCAAA 601
 QY 1067 ACTAGGAGC 1075
 Db 602 ACCAGGAGC 610

RESULT 3
 AAF15981
 ID AAF15981 standard; cDNA; 1113 BP.

AAF15981;

13-MAR-2001 (first entry)

Human prostate cancer antigen nucleotide sequence SEQ ID NO:416.

Human; prostate cancer; prostate cancer antigen; detection; diagnosis;

KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease; ss.

OS Homo sapiens.

PN W0200055174-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US05988.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.

PA (ROSE/) ROSEN C A.

XX Rosen CA, Ruben SM;

XX WPI; 2000-587513/55.

DR P-PSDB; AAB56778.

PT prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -

PS Claim 1; Page 900-901; 2338pp; English.

XX AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.

XX Sequence 1113 BP; 414 A; 189 C; 319 G; 190 T; 1 other;

XX Query Match

21.8%; Score 393; DB 21; Length 1113;

XX Best Local Similarity 96.4%; Pred. No. 3e-77; Mismatches 15; Indels 0; Gaps 0;

Matches 402; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 1 ACAGACAGAACTGCGCTTTTGGACAGAAAGTGGCTTGGCGAGCTGCAAGAGATG 60
DB 658 ACAGACAGAACTGCGCTTTTGGACAGAAAGTGGCTTGGCGAGCTGCAAGAGATG 717
OY 61 GAATGACATTAATGCCAAATAATTAATATATAGAACTAGACAAAGCCCATTTGTT 120
DB 718 GAATGACATTAATGCCAAATAATTAATATATAGAACTAGACAAAGCCCATTTGTT 777
OY 121 TTATATCTGGAAGATGTGTCCAGTACCAAAAATAATAGAGAGTACAGAGAAA 180
DB 778 TTATATCTGGAAGATGTGTCCAGTACCAAAAATAATAGAGAGTACAGAGAAA 837
OY 181 AATGAAGCTTTATTTGAAGGTAGACGATGAAATTTGCAGAACAAATAATAATGA 240
DB 838 AATGAAGCTTTATTTGAAGGTAGACGATGAAATTTGCAGAACAAATAATAATGA 897
OY 241 GGCTAGGCTTAGAAGCAATCAATGAAGAAAAGACATCAGGTGCGCTAATGAAGA 300
DB 898 GGCTAGGCTTAGAAGCAATCAATGAAGAAAAGACATCAGGTGCGCTAATGAAGA 957
OY 301 ACAGAGGCGGACAAAGAGAGGTAGGTGCTCAGGAGAGAGAGTGGAGAGAGAC 360
|||||

DB 958 ACAGAGGCGGACAAAGAGAGGTAGGTGCTCAGGAGAGAGAGAGTGGAGAGAC 1017
OY 361 AGGTATCAGCACAATGATGTAGAAATAGAGACAGAGAGAGAGAGAAAGGA 417
DB 1018 AGGTATCAGCACAATGATGTAGAAAGAGAGAGAGAGAGAGAGAGAAAGGA 1074
|||||

RESULT 4

AAV89627 ID AAV89627 standard; cDNA; 325 BP.

XX AAV89627;

DT 15-FEB-1999 (first entry)

DE EST clone C0559.

XX Human; secreted protein; expressed sequence tag; EST; haematopoiesis;
KW tissue growth; activin; inhibin; chemotaxis; chemokines; haemostatic;
KW receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;
KW gene therapy; ss.

XX Homo sapiens.

XX W09845436-A2.

PN 15-OCT-1998.

PD 10-APR-1998; 98WO-US06955.

PF 10-APR-1997; 97US-0838821.

PR (GEM) GENETICS INST INC.

XX Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
PI Racie LA, Spaulding V, Treacy M;

XX WPI; 1999-070077/06.

DR New polynucleotides encoding human secreted proteins - derived from
XX e.g. human blood, kidney, foetal lung, placenta, testes, brain,
PT ovary, pituitary, retina and colon cDNA libraries.

XX Claim 1; Page 278; 618pp; English.

XX The present sequence represents a human expressed sequence tag (EST).
CC The polynucleotide, which is a secreted EST, and the encoded protein
CC are predicted to have useful biological activities which would make
CC them suitable for treating, preventing or ameliorating medical
CC conditions in humans and animals, although no supporting data is
CC given. Suggested activities include nutritional activity, immune
CC stimulating or suppressing activity, haematopoiesis regulating
CC activity, tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
CC activity, receptor/ligand activity, anti-inflammatory activity,
CC cadherin/tumour invasion suppressor activity, tumour inhibition
CC activity. The polynucleotide may also be useful for gene therapy.

XX Sequence 325 BP; 123 A; 47 C; 87 G; 68 T; 0 other;

XX Query Match

16.6%; Score 299.4; DB 20; Length 325;

XX Best Local Similarity 99.7%; Pred. No. 9.3e-57; Mismatches 1; Indels 0; Gaps 0;

Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 504 CTGAGTAAAGGAAAGTGAGAGACAGATAGTCAAGCTGAAGAGTTATGATGTC 563
DB 18 CTAGGTAAAGGAAAGTGAGAGACAGATAGTCAAGCTGAAGAGTTATGATGTC 77
OY 564 TAGAGATGTTGAGATGTCAACATGTAATGCTGACCAGAGAGTAAATGAACATATC 623
DB 78 TAGAGATGTTGAGATGTCAACATGTAATGCTGACCAGAGAGTAAATGAACATATC 137
OY 624 GAGTGAAGGTAGAACCTTCAGAAAATGAAGCTAGCAAGAAATTTGAACCAATG 683
|||||

|||||
Db 138 GAGTGAAGCTAGAACCTTCAGAAAATGACCTAGCAAGATGGAACCAAGATGG 197
OY 684 AATTGAATGACCCAGATTAAGAATGTAATCCCTTCTCCTGGGAAGAGATGTCA 743
Db 198 AATTGAATGACCCAGATTAAGAATGTAATCCCTTCTCCTGGGAAGAGATGTCA 257
OY 744 GTGCTTAGACATGAGAAAGAGCTGAGAGAAAAGAGAAAGATCTGAGCCCCAAC 803
Db 258 GTGCTTAGACATGAGAAAGAGCTGAGAGAAAAGAGAAAGATCTGAGCCCCAAC 317
OY 804 C 804
Db 318 C 318

RESULT 5
AAT20032
ID AAT20032 standard; cDNA to mRNA; 293 BP.
XX AAT20032;
DT 17-JUL-1996 (first entry)

XX Human gene signature HUMGS01171.
XX
XX Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
XX Homo sapiens.
XX
XX WO9514772-A1.
XX
XX 01-JUN-1995.
XX
XX 11-NOV-1994; 94WO-JP01916.
XX
XX 12-NOV-1993; 93JP-0355504.
XX
XX (MATS/) MATSUBARA K.
XX (OKUB/) OKUBO K.
XX
XX Matsubara K, Okubo K;
XX
XX WPI; 1995-206931/27.
XX
XX
XX Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
XX
XX Claim 1; Page 542; 2245pp; Japanese.

XX A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in AAT19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.

XX Sequence 293 BP; 103 A; 38 C; 57 G; 80 T; 15 other;
SQ
Query Match 14.9%; Score 268.4; DB 16; Length 293;

Best Local Similarity 93.2%; Pred. No. 6.3e-50;
Matches 272; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
OY 1494 GATCTTCAAGAGTGAAGAGAGCCCGAAAATCAGACAGGAAGACAAAGCGTTAATGA 1553
Db 1 GATCTTCAAGAGTGAAG 60
OY 1554 AGAAGCCAGGCTTTCTTAGCCATCTTTGACAGAGAGATTTCTTGATTAAGAGATTA 1613
Db 61 AGAAGCCAGGCTTTCTTAGCCATCTTTGACAGAGAGATTTCTTGATTAAGAGATTA 120
OY 1614 CCTTCCCTGTAAGAGAGATGCTGCTTAAGATTCATGTTGTAAGAGATTTTGA 1673
Db 121 CCTTCCCTGTAAGAGAGATGCTGCTTAAGATTCATGTTGTAAGAGATTTTGA 180
OY 1674 AAATACAGACTGTTGTTTACAGACATCTTGTACTTTTGTCAATTTTGAAGATT 1733
Db 181 AAATACAGACTGTTGTTTACAGACATCTTGTACTGNTGATATNNNGTAAAGATT 240
OY 1734 ATTTATCAAAATTAATGAGGTTCCAAAATATGTAATAATGATAATATAA 1785
Db 241 ATTTATCAAAATTAATGAGGTTCCAAAATATGTAATAATGATAATATAA 292

RESULT 6
AAC02581
ID AAC02581 standard; cDNA; 264 BP.
XX AAC02581;
AC
XX 06-OCT-2000 (first entry)
DT
XX
XX Human secreted protein 5' EST, SEQ ID NO: 2579.
DE
XX
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens.
OS
XX EPI033401-A2.
PN
XX 06-SEP-2000.
PD
XX 21-FEB-2000; 2000EP-0200610.
PF
XX 26-FEB-1999; 99US-0122487.
PR
XX
XX (GEST) GENSET.
PA
XX Dumas Milne Edwards J, Duclert A, Giordano J;
PI
XX
XX WPI; 2000-500381/45.
DR
XX P-PSDB; AAG02575.
DR
XX

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
PT
XX
XX Claim 1; SEQ ID 2579; 71pp + CD-ROM; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC cDNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion

CC vectors.
XX Sequence 264 BP; 80 A; 73 C; 53 G; 58 T; 0 other;
SQ
Query Match 14.6%; Score 262.4; DB 21; Length 264;
Best Local Similarity 99.6%; Pred. No. 1.3e-48;
Matches 263; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 677 GAAATGGAATTGAATGAGCCAGATTAAGATGTAATCCCTTCTCTGGGAAAGAG 736
DB 1 GAAATGGAATTGAATGAGCCAGATTAAGATGTAATCCCTTCTCTGGGAAAGAG 60
OY 737 AATGTCAGTCTTAGACATGAGAAAGAGCTGAGAGAAAGAGAAAGAAATCTGAG 796
DB 61 AATGTCAGTCTTAGACATGAGAAAGAGCTGAGAGAGAAAGAGAAAGAAATCTGAG 120
OY 797 CCCCACCTGAGCTGTGGCTCAACCTCAGCTCAGCTCAGCCAGCTCAGCTTCAA 856
DB 121 CCCCACCTGAGCTGTGGCTCAACCTCAGCTCAGCTCAGCCAGCTCAGCTTCAA 180
DB 857 TCCAGTCCCAACAGTACTCCAGTCCAGCTCCCTCTCAGCTGAGATTTGTCATTA 916
DB 181 TCCAGTCCCAACAGTACTCCAGTCCAGCTCCCTCTCAGCTGAGATTTGTCATTA 240
OY 917 GCTGTTTACAGCCCAACCCCAA 940
DB 241 GCTGTTTACAGCCCAACCCCAA 264

SID 692

RESULT 7
AAV86711
ID AAV86711 standard; cDNA; 263 BP.
XX
AC AAV86711;
XX
DT 27-APR-1999 (first entry)
XX
DE EST clone AW32.

XX Expressed sequence tag; secreted protein; haematopoiesis regulator;
KW tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
XX

OS Homo sapiens.
XX
PN W09845435-A2.
XX
15-OCT-1998.

PF 10-APR-1998; 98WO-US06954.
XX
PR 10-APR-1997; 97US-0835913.
XX

PA (GEMT) GENETICS INST INC.

PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
PI Racie LA, Spaulding V, Treacy M;

DR WPI; 1999-070076/06.

XX New polynucleotides encoding human secreted proteins - derived from
PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
PT ovary, pituitary, retina and colon cDNA libraries

PS Claim 1; Page 324; 633pp; English.

CC This sequence represents an expressed sequence tag (EST), and is a
CC polynucleotide of the invention. The polynucleotides of the invention are
CC all secreted EST sequences isolated from a variety of human tissue
CC sources. The EST sequences and proteins encoded by them are predicted to
CC have useful biological activities which would make them suitable for
CC treating, preventing or ameliorating medical conditions in humans and

CC animals, although no supporting data is given. Suggested activities
CC include nutritional activity, immune stimulating or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
CC therapy. The EST sequences are also stated to be useful for gene
CC

XX Sequence 263 BP; 107 A; 38 C; 76 G; 42 T; 0 other;

SQ
Query Match 14.0%; Score 251.4; DB 20; Length 263;
Best Local Similarity 99.6%; Pred. No. 3.5e-46;
Matches 252; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 127 TCCTGGAAGATGTGTCCAGCTAACCAAACTAATAGAGAGTCACAGAGAAATGAA 186
DB 2 TCCTGGAAGATGTGTCCAGCAACCAAACTAATAGAGAGTCACAGAGAAATGAA 61
OY 187 CGCTTATTTGAGGTAGACGCAATCGAATTTGACAGAACAAATTAATGAGGCTAG 246
DB 62 CGCTTATTTGAGGTAGACGCAATCGAATTTGACAGAACAAATTAATGAGGCTAG 121
OY 247 GCCTAGAACCAATCAATGAAGAGAAAGAGATCAGGTGCTGCTAATGAGAGACGAA 306
DB 122 GCCTAGAACCAATCAATGAAGAGAAAGAGATCAGGTGCTGCTAATGAGAGACGAA 181
OY 307 GCGGAAACAAGAGAGGTAAGTGTGCTCAGCAGAGAGAGAGAGAGAGAGAGTAA 366
DB 182 GCGGAAACAAGAGAGGTAAGTGTGCTCAGCAGAGAGAGAGAGAGAGAGAGTAA 241
OY 367 TCAGCACCAATGAT 379
DB 242 TCAGCACCAATGAT 254

RESULT 8
AAS24061
ID AAS24061 standard; cDNA; 239 BP.

XX
AC AAS24061;

XX
DT 07-NOV-2001 (first entry)

XX Human ovarian PCR-subtracted cDNA library clone #242.

XX Immunogenic protein; cancer; ovarian tumour; T-cell stimulation; ss;
KW gene therapy; cytostatic; T-cell expansion; nucleic acid hybridisation;
KW primer; probe.

XX Homo sapiens.

OS
XX W0200157207-A2.

PN
XX 09-AUG-2001.

PD
XX 05-FEB-2001; 2001WO-US03733.

PF
XX 04-FEB-2000; 2000US-0180403.

PR
XX 28-MAR-2000; 2000US-0192745.

XX (CORI-) CORIXA CORP.

PA
XX Algate PA, Mannion J;

PI
XX WPI; 2001-488879/53.

XX New polynucleotides encoding ovarian tumour proteins, useful for
PT treating ovarian cancer, and as probes, primers, and markers of cancer
PT progression

XX Example 1; page 149; 378pp; English.

CC The invention comprises compositions used for the therapy and diagnosis
CC of ovarian cancer. The compositions comprise one or more ovarian tumour
CC proteins, their associated polynucleotides, or immunogenic portions of
CC the proteins. The ovarian tumour polynucleotides and polypeptides are
CC useful for stimulating and/or expanding T cells specific for a tumour
CC protein. They are also useful for inhibiting the development of cancer in
CC a patient with an ovarian tumour DNA or protein by incubating isolated
CC T-cells allowing them to proliferate, and administering to the patient.
CC The sequences can be used as markers for cancer, for example, to monitor
CC ovarian cancer progression. Probes and primers are useful in nucleic acid
CC hybridisation, in detecting the presence of complementary sequences in a
CC given sample, for preparing mutant species and for preparing other
CC genetic constructions. Sequences AAS23820-AAS25231 and AAS25228-AAS25549
CC represent human ovarian tumour protein CDNA clones.

XX Sequence 239 BP; 81 A; 28 C; 74 G; 42 T; 14 other;

Query Match 9.2%; Score 166.4; DB 22; Length 239;
Best Local Similarity 93.4%; Pred. No. 2.1e-27;
Matches 170; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

1203 GTCGCAGTAGTTCAGTAGCAGCTCCAGTACAAGTGGCAGCAGCAGAGATAGTAGCA 1262
|||||
1 GTCGCAGTAGTTCAGTAGCAGCTCCAGTACAAGTGGCAGCAGCAGAGATAGTAGCA 60

OY 1263 GTAGCACTAGTAGTAGTAGTAGAGTAGAGTCCGAGTAGGGTCCGGGACATAATAGAG 1322
|||||
61 GTAGCACTAGTAGTAGTAGTAGAGTAGAGTCCGAGTAGGGTCCGGGACATAATAGAG 120

OY 1323 ATAGAAAGCACAGAGAGCGCTGCATCGGAAGAGAGGATATTCTCAGGACTAGAAAGAA 1382
|||||
121 ATAGAAAGCACAGAGAGCGCTGCATCGGAAGAGAGGATATTCTCAGGACTAGAAAGAT 180

OY 1383 GT 1384
||
Db 181 GT 182

RESULT 9
AAH82631

ID AAH82631 standard; cDNA; 239 BP.

AC AAH82631;

DT 25-SEP-2001 (first entry)

DE Human ovarian tumour associated polynucleotide sequence SEQ ID NO:255.

XX Human; ovarian tumour; ovarian cancer; diagnosis; gene therapy;

XX Immunogenic; vaccine; ss.

OS Homo sapiens.

XX WO200151513-A2.

PD 19-JUL-2001.

PF 16-JAN-2001; 2001WO-US01575.

PR 14-JAN-2000; 2000US-0176722.

PA (CORI-) CORIXA CORP.

PI Algate PA;

DR MPI; 2001-425866/45.

XX Novel ovarian tumor proteins, and nucleic acids encoding them, used to
PT treat and diagnose cancers, particularly ovarian cancer -
PT Claim 5; Page 109; 338pp; English.

XX AAH82377 to AAH83878 represent human ovarian tumour-associated

CC polynucleotide sequences which encode ovarian tumour proteins. The
CC ovarian tumour protein and polynucleotide sequences have cytostatic
CC activity, and can be used in gene therapy and vaccine production. The
CC ovarian tumour proteins and polynucleotides can be used to inhibit
CC the development of cancer, particularly ovarian cancer. They can also
CC be used to diagnose the onset and progression of cancer.

XX Sequence 239 BP; 81 A; 28 C; 74 G; 42 T; 14 other;

Query Match 9.2%; Score 166.4; DB 22; Length 239;
Best Local Similarity 93.4%; Pred. No. 2.1e-27;
Matches 170; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

OY 1203 GTCGCAGTAGTTCAGTAGCAGCTCCAGTACAAGTGGCAGCAGCAGAGATAGTAGCA 1262
|||||
1 GTCGCAGTAGTTCAGTAGCAGCTCCAGTACAAGTGGCAGCAGCAGAGATAGTAGCA 60

OY 1263 GTAGCACTAGTAGTAGTAGTAGAGTAGAGTCCGAGTAGGGTCCGGGACATAATAGAG 1322
|||||
61 GTAGCACTAGTAGTAGTAGTAGAGTAGAGTCCGAGTAGGGTCCGGGACATAATAGAG 120

OY 1323 ATAGAAAGCACAGAGAGCGCTGCATCGGAAGAGAGGATATTCTCAGGACTAGAAAGAA 1382
|||||
121 ATAGAAAGCACAGAGAGCGCTGCATCGGAAGAGAGGATATTCTCAGGACTAGAAAGAT 180

OY 1383 GT 1384
||
Db 181 GT 182

RESULT 10
AAT05868

ID AAT05868 standard; DNA; 3399 BP.

AC AAT05868;

DT 14-AUG-1996 (first entry)

DE Chicken leucocytozoan DNA encoding immunogenic protein for vaccines.

XX Chicken leucocytozoan; Immunogen; recombinant vaccine; protection;

XX Immunisation; vaccination; ss.

OS Chicken leucocytozoan.

XX Key Location/Qualifiers

FT CDS 1..3399

FT misc_feature 1150..3218

FT /note- b
FT use as insert in a recombinant vaccine
FT against chicken leucocytozoan disease

PN JP07284392-A.

PD 31-OCT-1995.

PF 19-APR-1994; 94JP-0080643.

PR 19-APR-1994; 94JP-0080643.

PA (DOBU-) DOBUTSUO SEIBUTSUGARUTEKI SEIZAI KYOKAI.

PA (KITA) KITASATO KENKYUSHO SH.

DR MPI; 1996-006311/01.

DR P-PSDB; AAR97866.

XX Chicken leucocytozoan immunogenic protein - used in a recombinant
PT vaccine against chicken leucocytozoan disease
PT Claim 6; Page 6-9; 35pp; Japanese.

DR P-PSDB; AAY96255.

XX Treating or preventing a disease associated with rhodino virus
PT Infection in a mammal which includes Kaposi's Sarcoma and Primary
PT Efusion Lymphoma
XX
PS Disclosure; Fig 6; 70pp; English.

CC The present sequence is the Kaposi's sarcoma-associated herpesvirus,
CC (KSHV) latency-associated nuclear antigen (LANA) gene. KSHV is also known
CC as Human Herpes Virus 8 (HHV8) and belongs to the rhadino virus, or
CC gamma-2 herpes virus class. The LANA protein is necessary for the
CC efficient persistence of rhadino virus DNA in mammalian cells. Persistent
CC rhadino virus infection is implicated in a variety of diseases e.g.
CC Kaposi's Sarcoma (KS), Primary Efusion Lymphoma (PEL) and multicentric
CC Castleman's disease. In addition, KS is a common malignancy in HIV
CC patients. KSHV persists in host cells in a latent form. One of the few
CC genes expressed from the latent viral DNA is LANA. LANA associates with
CC both human chromosomes and with the rhadino virus cis-acting element
CC (RVCAE), thereby providing a tethering function: the KSHV DNA episome is
CC "tied" to the host chromosomes. This allows the viral DNA to persist in
CC the host cell. The present sequence may be used to screen and identify
CC molecules that inhibit LANA interaction with RVCAE, thereby interfering
CC with the latency cycle of this virus. Potential antiviral treatments for
CC the above mentioned diseases may therefore be based on LANA deregulation.

SQ Sequence 3489 BP; 1053 A; 862 C; 1137 G; 437 T; 0 other;

Query Match 6.7%; Score 121.2; DB 21; Length 3489;
Best Local Similarity 45.9%; Pred. No. 5.4e-17;
Matches 492; Conservative 0; Mismatches 573; Indels 7; Gaps 2;

OY 293 AATGAGACAGAGAGCGGACAAAGAGAGGCTAAGGTGCTCAGCAGAGAGAGAGTGTG 352
DB 982 AATAAGATGATGAGAGAGAGAGAGAGATGAGAGAGAGAGAGAGATGAGAGAGAG 1041
OY 353 GAGGAGACAGTATATCAGCACAATGATGTAAGAAATAGAGAGAGAGAGAGAGAGAA 412
DB 1042 GATGACGAGAGATGACGAGAGAGATGACGAGAGAGATGACGAGAGATGACGAGAG 1101
OY 413 AAGCAATAGCGATGTTTCATAGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 472
DB 1102 GATGACGAGAGAGATGACGAGAGAGATGACGAGAGAGATGACGAGAGAGAGAGAG 1161
OY 473 CAGCAATGAGGTTAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 532
DB 1162 GAGGAT 1221
OY 533 GATAGTCAGCCTGAGAGAGTATGATGCTAGAGATGCTGAGAGAGAGAGAGAGAGAGTA 592
DB 1222 GATGAGAGAGATGAG 1281
OY 593 ATTGCTGACGAGAGGTAATGAACTAATCGAGTTGAAAGTGTAGAACCTTCAGAAAT 652
DB 1282 GACGGGGGCG--GATGAAACMAAACGTTGAGCATCCAAAGTTCCACACAGCAGCAGAG 1338
OY 653 GAAGCTAGCAAGAATTGGAACCAAGAAATTTGAATTGAGCCAGATTAAGAATGT 712
DB 1339 CCACAAACAGCA 1398
OY 713 AATGCCCTTCTCTGGAAGAAGATGTCAGTCTTAGACATGAAAGAGAGTCTGAG 772
DB 1399 CAACAGCA 1458
OY 773 GAAAAAGAGAAAAAGATCTGAGCCCACTGAGCCTGTGCTCAACCTCAGCCTCAG 832
DB 1459 CAGCAG 1518
OY 833 TCTCAG 892
DB 1519 CAGGAG 1578
OY 893 TCTCAG 952

DB 1579 CAGGAG 1638
OY 953 CAGGAG 1012
DB 1639 CAGGAG 1698
OY 1013 GTACAGTAG 1068
DB 1699 CCGGAG 1758
OY 1069 TAGGAG 1128
DB 1759 CCGGAG 1818
OY 1129 TAGCAGTAG 1188
DB 1819 CAGGAG 1878
OY 1189 TGAAGTAG 1248
DB 1879 CAGGAG 1938
OY 1249 CAGGAG 1308
DB 1939 CAGGAG 1998
OY 1309 GGGACATATAGAGATAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1360
DB 1999 CAGGAG 2050
RESULT 12
AAF82901
ID AAF82901 standard; DNA; 3489 BP.
XX
AC AAF82901;
XX
DT 29-JUN-2001 (first entry)
XX
DE Nucleotide sequence of KSHV tethering protein, LANA.
XX
KW Histone H1; tethering protein; LANA; gene therapy; multiple sclerosis;
KW Parkinson's disease; Huntington disease; diabetes; human herpesvirus 8;
KW KSHV; latency-associated nuclear antigen; LANA; ds.
XX
OS Kaposi's sarcoma associated herpesvirus.
XX
FH Key Location/Qualifiers
FT CDS 1..3489
FT /tag- a
XX
PN WO200125484-A2.
XX
XX 12-APR-2001.
PD
XX 29-SEP-2000; 2000WO-US26908.
PF
XX 01-OCT-1999; 99US-0410399.
PR
XX (UNMI) UNIV MICHIGAN.
PA
PI Robertson ES, Cotter MA;
XX
DR WPI: 2001-281736/29.
XX P-PSDB; AAB62331.
XX
PT A composition for use in gene therapy comprises an expression vector
PT that includes a nucleic acid sequence encoding a nucleic acid binding
PT protein -
XX
PS Disclosure; Fig 9A; 60pp; English.
XX

XX Sequence 32207 BP; 7229 A; 9156 C; 8713 G; 7109 T; 0 other;
SO

Query Match	6.78;	Score 121.2;	DB 20;	Length 32207;
Best Local Similarity	45.98;	Pred. No. 1.2e-16;		
Matches 492;	Conservative	0;	Mismatches 573;	Indels 7;
				Gaps 2;

OY	293	ATGAAAGAACAGAAAGCGCGAACAAGAGGGGTAAAGTGGCTCAGCGAGAGAAAGTTG	352
Db	21015	AAATGAAGATGATGAGGAGAGCAGAGACAGATGAGGAGGACGAGAGATGACGAGAG	20956
OY	353	GAGGAGACAGGTAATCAGCACAAATGATGTAGAAATAGAGGAAGCAGAGAGAGAGAA	412
Db	20955	GATGACGAGGAGGATGACGAGGAGATGACGAGGAGGATGACGAGAGAGATGACGAGGAG	20896
OY	413	AAGGAATAGCGATTGTTTCATAGTGTATGACAGAAAGAACAGAGAGAGAAACAATAA	472
Db	20895	GATGACGAGGAGGATGACGAGGAGATGACGAGGAGGATGACGAGGAGATGACGAGGAG	20836
OY	473	CAGGAATGAGGTTAAGATGAGAGAGAACTGAGGTAAGGGAAGTGAGAACGACGAG	532
Db	20835	GAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT	20776
OY	533	GATAGTCAGCCTGAAGAAGTTATGATGTGCTAGAGATGTTGAGAAATGTCAAATGTA	592
Db	20775	GATGAGCAATGAGGAGGAGGAGGATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	20716
OY	593	ATTGCTGACGAGGAGGTAATGMAACTAAATCGAGTTGAAAGTGTAGAACTTCAGAAAT	652
Db	20715	GAGCGGGGC---GATGGAACAATAACGTTGAGCATCCAAAGTTCACAACAGCAGCAGGAG	20659
OY	653	GAACTAGCAAAAGAAATTGGAACCAAGAAATGGAATTTGAAATTGAGCCAGATAAAGAAATGT	712
Db	20658	CCACAACAGCAGGAGGCCACAGCAGCAGAGCCACAGCAGCAGAGAGCCCTGCAGAGCCCA	20599
OY	713	AAATCCCTTCTCCTGGGAAAGAGATGTCACTGCTTTAGACATGGAAGGAGCTGTAG	772
Db	20598	CAACAGCAGGAGCCACAGCAGCAGGAGCCACAGCAGCAGGAGGCCCTGCAGAGCCACAA	20539
OY	773	GAAGAAAGAGAAAAAGAAATCTGAGCCCCCAACTGAGCCTGTGGCTCAACCTCAGCCTCAG	832
Db	20538	CAGCAGAGCCACAGCAGCAGGAGCCCTGACAGGAGCCACAACAGCAGGAGCCACAACAG	20479
OY	833	TCCTAGCCCCAGCTTCACTCAATCCAGTCCCAACAGTACTCCAGTCCAGCCTCCC	892
Db	20478	CAGGAGCCACAGCAGCAGGAGCCACAGCAGCAGGAGCCACAGCAGCAGGAGCCACAGCAG	20419
OY	893	TCCTAGCCTGAGGATTTGTCTATAGCTGTTTACAGCCACACCCCAAGTTACTCAGGAG	952
Db	20418	CAGGAGCCACAGCAGCAGGAGGCCACAGCAGCAGGAGCCACAGCAGGAGCCACAGCAG	20359
OY	953	CAAGGCGATTACTACCTGAGAGGAAGATTTCCTGTAGACTGTAAACTCTACTGAG	1012
Db	20358	CAGGAGCCACAGCAGCGGAGCCACAGCAGCGGAGGCCACAGCAGCGGAGCCACAGCAG	20299
OY	1013	GTACCACTAGAGCCAGTCTTGACAGTACATCCAGAG---AGCAAGAGCAAAAACCAAAAC	1068
Db	20298	CGGAGCCACAGCAGCGGAGCCACAGCAGCGGAGCCACAGCAGCGGAGCCACAGCAG	20239
OY	1069	TAGAGCAGAGTAGAGTGCAGCTAGAAATAAACAAGCAAGAGTAGAAAGTGAAGTGAAGCAG	1128
Db	20238	CGGAGCCACAGCAGCGGAGCCACAGCAGCAGGATGACAGCAGCAGGATGACGAGCAG	20179
OY	1129	TAGCAGTAGCAGTTCTAGTAGCAGTTCAACCACTAGCAGCAGTGAAGTAGTCCAGCAG	1188
Db	20178	CAGGATGACAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAG	20119
OY	1189	TGGAAGTAGTAGCAGTCCGAGTAGTTCACAGTAGCAGCTCCAGTACAAAGTGGCAGCAGCAG	1248
Db	20118	CAGGATGACAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAG	20059
OY	1249	CAGAGATAGTAGCAGTAGCACTAGTAGTAGTAGAGTAGAAGTCGAGTAGGGTGC	1308

Db	20058	CAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGCAGGATGAGCAGCAG	19999
QY	1309	GGGACATATAGAGATAGAAAGCACAGAGAGCGCTGCATCGGAAGAGAAG	1360
Db	19998	CAGCATGACACAGCAGCAGGATGAGCAGCAGCAGCAGATGAGCAGCAGCAGCAGATG	19947

ID	AAV19941/c	AAV19941 standard; DNA; 137507 BP.
XX	AC	AAV19941;
XX	DT	03-AUG-1998 (first entry)
XX	DE	KSHV long unique coding region and terminal repeat.
KM	KSHV; HHV8; human herpes virus 8; macrophage inflammatory protein 1B; interleukin-6; IL-6; interferon regulatory factor; rheumatoid complement-binding protein; glycoprotein B; capsid protein IV; immediate early protein; Kaposi's sarcoma; protective vaccine lymphoproliferative disease; leukaemia; splenomegaly; mycosis HIV immune status; anti-inflammatory agent; therapy; ds.	
OS	Kaposi's sarcoma-associated herpes virus.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1142..2794
FT		/tag- a
FT		/product- complement-binding protein
FT	CDS	8699..11236
FT		/tag- b
FT		/product- glycoprotein B
FT	CDS	complement (17261..17875)
FT		/tag- c
FT		/product- interleukin 6
FT	CDS	complement (21548..21832)
FT		/tag- d
FT		/product- macrophage inflammatory protein II
FT	CDS	complement (27137..27424)
FT		/tag- e
FT		/product- interferon regulatory factor 1
FT	CDS	28661..29741
FT		/tag- f
FT		/product- protein T1.1
FT	CDS	complement (58976..60175)
FT		/tag- g
FT		/product- glycoprotein M
FT	CDS	complement (69412..69915)
FT		/tag- h
FT		/product- glycoprotein L
FT	CDS	complement (88410..88910)
FT		/tag- i
FT		/product- interferon regulatory factor 2
FT	CDS	89600..90541
FT		/tag- j
FT		/product- interferon regulatory factor 3
FT	CDS	90173..90643
FT		/tag- k
FT		/product- glycoprotein X
FT	CDS	complement (93636..94127)
FT		/tag- l
FT		/product- interferon regulatory factor 4
FT	CDS	complement (111931..112443)
FT		/tag- m
FT		/product- capsid protein IV
FT	CDS	complement (123808..127296)
FT		/tag- n
FT		/product- immediate early protein
PN	WO9804576-A1.	
XX		
PD	05-FEB-1998.	

XX 22-JUL-1997; 97MO-US13346.
 XX 29-NOV-1996; 96US-0757669.
 PR 25-JUL-1996; 96US-0686243.
 PR 25-JUL-1996; 96US-0686349.
 PR 25-JUL-1996; 96US-0686350.
 PR 25-JUL-1996; 96US-0687253.
 PR 25-JUL-1996; 96US-0688814.
 PR 05-SEP-1996; 96US-0708678.
 PR 10-OCT-1996; 96US-0728323.
 PR 13-NOV-1996; 96US-0747887.
 PR 13-NOV-1996; 96US-0748640.
 XX (UNCO) UNIV COLUMBIA NEW YORK.
 PI Bohenzky RA, Chang Y, Edelman IS, Moore PS, Russo JJ;
 DR WPI; 1998-130615/12.
 XX New nucleic acid encoding Kaposi's sarcoma associated herpes virus
 proteins - useful for, e.g. detecting levels of HHV8 in, and
 PT preparation of vaccines for treatment of, HIV patients
 XX
 PS Example 2; Page 135-203; 230pp; English.
 XX
 CC This sequence represents the long unique region and terminal repeat of
 CC the Kaposi's sarcoma-associated herpes virus (KSHV). KSHV is also known
 CC as human herpes virus 8 (HHV8). This sequence contains the DNAs of the
 CC invention which encode KSHV polypeptides selected from: (a) viral
 CC macrophage inflammatory protein (MIP) II; (b) viral interleukin-6 (IL-6);
 CC (c) viral IRF 1; (d) complement-binding protein; glycoproteins B, M or L;
 CC (d) capsid protein IV encoded by ORF65; and (e) immediate early protein
 CC encoded by ORF73. Labelled probes for the nucleic acid, proteins encoded
 CC by it, and antibodies (Ab) specific for the nucleic acid, proteins encoded
 CC detecting HHV8, specifically for diagnosis of Kaposi's sarcoma, in body
 CC fluids or tissue samples. HHV8 infections can be treated with antisense
 CC or triplex forming molecules or agents that bind specifically to the
 CC protein. Ab may be used for prophylaxis or treatment of HHV8 infection,
 CC while the protein can be used in protective vaccines. Ab may also be used
 CC to differentiate between lymphomas, and HHV8 may be implicated in many
 CC other lymphoproliferative diseases such as lymphomas, leukaemia,
 CC splenomegaly and mycosis fungoides. Cells and animals containing the
 CC nucleic acid are useful for drug screening. HHV8-derived peptides can be
 CC used as targets for antiviral drugs, e.g. dihydrofolate reductase gene
 CC can be inhibited with methotrexate. These can also be used to determine
 CC the immune status of a patient infected with HIV. HHV8 derived protein
 CC viral MIP III may be used as an anti-inflammatory agent for,
 CC e.g. treating rheumatoid arthritis. This sequence is stated as containing
 CC 81 open reading frames.
 XX
 SO Sequence 137507 BP; 32579 A; 37795 C; 35758 G; 31375 T; 0 other;
 Query Match 6.7%; Score 121.2; DB 19; Length 137507;
 Best Local Similarity 45.9%; Pred. No. 2.1e-16;
 Matches 492; Conservative 0; Mismatches 573; Indels 7; Gaps 2;
 OY 293 AATGAGACAGAGAGCGGGAACAAGAGAGGTGCTCAGCAGAGAGAGAGTTG 352
 DB 126315 AATGAGATGATGAGAGAGAGAGATGATGAGAGAGAGATGATGAGAGAG 126256
 OY 353 GAGGAGACAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 412
 DB 126255 GATGACGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 126196
 OY 413 AAGGAATAGCATTTTCATAGTATGATGATGATGATGATGATGATGATGATGATGAT 472
 DB 126195 GATGACGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 126136
 OY 473 CAGGAATGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 532
 DB 126135 GAGGACGATGATG 126076

OY 533 GATAGTACAGCTGAAGAAGTTATGATGCTAGAGATGCTTGAAGATGTCAAACATGTA 592
 DB 126075 GATGAGACAAATAGAGACGAGAGAGATGATGAGAGAGAGAGAGAGAGAGAGAG 126016
 OY 593 ATTGCTGACGAGAGGTAAATGAACTAAATGAGTTGAAAGTTAGAACTTCAGAAAT 652
 DB 126015 GACGGGGC--GATGAAACAAACAGTTGAGCATCCAAAGTTCAACAACAGCAGCAG 125959
 OY 653 GAAGCTAGCAAGAATTTGAAACCAAGAAATGGAATTTGAAATTTGAGCCAGATTAAGAATGT 712
 DB 125958 CCATTAACAGCAGAGACCCACAGCAGCAGAGCCACAGCAGCAGAGCCCTGACAGAGCCA 125899
 OY 713 AATCCCTTTCTCTGGGAAAGAAATGTCAGTCTTAGACATGGAAGAGAGTCTGAG 772
 DB 125898 CAACAGCAGAGACCAACAGCAGCAGAGCCACAGCAGAGAGCCCTGACAGAGCCACA 125839
 OY 773 GAAAAAGAGAAAGAAATCTGAGCCCAACCTGAGCCTGTGCTCAACCTCAGCCTCAG 832
 DB 125838 CAGCAGAGCCACAGCAGCAGAGAGCCCTGACAGAGCCACACAGCAGAGCCACACAG 125779
 OY 833 TCTACGCCAGCTTCAAGCTTCAATCCCAACCAACAGTACTCCAGCCTCC 892
 DB 125778 CAGGAGCCACAGCAGCAGAGAGCCACAGCAGCAGAGAGCCACAGCAGAGAGCCACAGCAG 125719
 OY 893 TCTACGCTGAGAGTTTGTCTATTAGCTGTTTACAGCCCAACCCCAAGTTACTCAGAG 952
 DB 125718 CAGGAGCCACAGCAGCAGAGAGCCACAGCAGCAGAGAGCCACAGCAGAGAGCCACAGCAG 125659
 OY 953 CAAGGCATTACTACCTGAGAGGAAGATTTCTGTAGAGTCTGTAATAACTCAGCTGAG 1012
 DB 125658 CAGGAGCCACAGCAGCAGAGAGCCACAGCAGCAGAGAGCCACAGCAGAGAGCCACAGCAG 125599
 OY 1013 GTACAGTAGAGCCAGTCTTGACAGTATCAGAG--AGCAAGAGCAAAACCAAAAC 1068
 DB 125598 CCGGAGCCACAGCAGCAGAGAGCCACAGCAGCAGAGAGCCACAGCAGAGAGCCACAGCAG 125539
 OY 1069 TAGGACAGAGTAGAGTCTGAGCTAGAAATTAACAAGCAAGAGTAGAAGTGAAGCAG 1128
 DB 125538 CCGGAGCCACAGCAGCAGAGAGCCACAGCAGCAGAGATGAGCAGCAGAGATGAGCAGCAG 125479
 OY 1129 TAGCAGTAGAGTCTTAGTAGCAGTTCAACAGTAGCAGCAGTGAAGTAGTCCAGCAG 1188
 DB 125478 CAGGATGAGCAGCAGCAGAGATGAGCAGCAGCAGAGATGAGCAGCAGAGATGAGCAGCAG 125419
 OY 1189 TGAAGTAGTAGAGTCTGAGTAGTTCAGTAGCAGTCCAGTACAAGTGGCAGCAGCAG 1248
 DB 125418 CAGGATGAGCAGCAGCAGAGATGAGCAGCAGCAGAGATGAGCAGCAGAGATGAGCAGCAG 125359
 OY 1249 CAGAGTAGTAGCAGTACACTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAG 1308
 DB 125358 CAGGATGAGCAGCAGCAGAGATGAGCAGCAGCAGAGATGAGCAGCAGAGATGAGCAGCAG 125299
 OY 1309 GGGACATATATAGATAGAAAGCAGAGAGAGCGCTGATCGGAGAGAGAG 1360
 DB 125298 CAGGATGAGCAGCAGCAGAGATGAGCAGCAGCAGATGAGCAGCAGAGATGAGCAGCAGATG 125247
 RESULT 15
 AAA59242/c
 ID AAA59242 standard; cDNA; 1037 BP.
 XX
 AC AAA59242;
 XX
 DT 07-NOV-2000 (first entry)
 XX
 DE Exons E, C and A of the SCAB repeat region.
 XX
 KW Repeat region; spinocerebellar ataxia type 8; SCAB; long arm;
 KW chromosome 13; polymorphic CTA repeat; CTG repeat; ss.
 XX
 OS Homo sapiens.
 XX
 PN CA2283758-A1.

XX 28-APR-2000.
PD

PF 27-SEP-1999: 99CA-2283758.

PR 28-OCT-1998; 9805-0181585.

PA (MINU) UNIV MINNESOTA.

PI Bezow KA, Moseley-alldredge ML, Ranum LPW, Koob MD;

DR WPI; 2000-491456/44.

PT Novel nucleic acid molecule containing repeat region of spinocerebellar
PT ataxia type 8 coding sequence, useful as primer or probe for diagnosing
PT or determining risk for developing spinocerebellar ataxia -

PS **Claim 25; Fig 7C; 68pp; English.**

The present sequence represents a nucleic acid molecule containing a repeat region of a spinocerebellar ataxia type 8 (SCA8) coding sequence which is located within the long arm of chromosome 13. The SCA8 coding sequence contains polymorphic CTA repeats and CTG repeats. Generally, the CTA repeat is unstable, and can vary between individuals in different families (the number of CTA repeats in the region does not vary between individuals of a family). Fragments of the nucleic acid molecule are useful as primers and probes for determining whether an individual has or is at risk of developing spinocerebellar ataxia 8.

SQ Sequence 1037 BP; 229 A; 261 C; 242 G; 305 T; 0 other;

Query Match	Score	DB	Length
6.68;	118;	21;	1037;

Matches 169; **Conservative** 0; **Mismatches** 85; **Indels** 0; **Gaps** 0;

1049 AGCAAGAGCAAAACCAAACTAGGAGCAGAAGTAGAGGTCGAGCTAGAAATAAACCAAGC 1108

[illegible]

1109 AAGAGTAGAAGTCGAAGCAGTAGCAGTCTCTAGTAGCAGTTCAACCCAGTAGCAGC 1168

[illegible]

QY 1169 AGTGAAGTAGTTCACGACGTGGAAGTAGTAGCAGTCCGACAGTAGCTCC 1228

[illegible]

1229 AGTACAAGTGGCAGCAGCAGAGATAGTAGCAGTAGCACTAGTAGTAGTGAGAGT 1288

711 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGTAGTAGTAGTAGT 652

QY 1289 AGAAGTCGGAGTAG 1302

Db 651 AGTACTAGTACTAG 638

Search completed: October 16, 2002, 13:05:14
Job time : 380 secs